LIQUIDTEXT: SUPPORTING ACTIVE READING THROUGH FLEXIBLE DOCUMENT REPRESENTATIONS

A Dissertation
Presented to
The Academic Faculty

by

Craig S Tashman

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Interactive Computing

Georgia Institute of Technology
May, 2012

LIQUIDTEXT: SUPPORTING ACTIVE READING THROUGH FLEXIBLE DOCUMENT REPRESENTATIONS

Approved by:

Dr. W. Keith Edwards, Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. John T. Stasko
School of Interactive Computing
Georgia Institute of Technology

Dr. Gregory D. Abowd
School of Interactive Computing
Georgia Institute of Technology

Dr. Ravin Balakrishnan
Department of Computer Science
University of Toronto

Dr. Elizabeth Mynatt
School of Interactive Computing
Georgia Institute of Technology

Date Approved: March 27, 2012

To Jesus Christ, my God and Savior.
If any wisdom is found in these pages, it is to His credit.

ACKNOWLEDGEMENTS

While a dissertation is fundamentally a solitary endeavor, it is likewise impossible without the effort and care of multitudes of individuals. The first of these is my advisor, Keith Edwards for indefatigably working to support my research. Through all the crazy inventions, rabbit trails, and touch screens, Keith has been relentlessly encouraging and enabling. A better advisor or mentor I could not ask for.

Equally important has been the wonderful group of undergraduate and master's students who worked with me: The interaction designers, Justin Ratcliff, Vignesh Swaminathan, Firaz Peer, Shiva KM, Xueli Xiao and Viraj Sapre who added brilliant insight into refining LiquidText. The study designers, Andy Echenique, Jason Roberts, and again Viraj Sapre who helped me devise and conduct the summative evaluation. And the data analysts, Pratik Zaveri, Chris Porter, Jade Sims, Megan Matthews, and Vishwaja Cheruku for their work in video coding.

I wish to thank the many who gave me expert feedback throughout the design process of LiquidText, but especially Jennifer Stoll and Yan Xu who offered tenacious encouragement, as well as excellent design criticism and ideas. Similarly, I wish to thank each of my committee members for taking the time to advise me on this research, as well as on preparing for life after graduation.

Likewise, I wish to thank Lloyd Treinish, Andy Conn, and Jim Wynne, for taking a big gamble on me a long time ago—it was instrumental in my getting to this point.

And especially, I wish to thank Marsha Tashman and David Tashman, for support and encouragement through many other crazy inventions. They are definitely the most awesome parents ever.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

Knowledge workers are frequently called upon to perform deep, critical reading involving a heightened level of interaction with the reading media and other tools. This process, known as *active reading*, entails highlighting, commenting upon, and flipping through a text, in addition to other actions. While paper is traditionally seen as the ideal medium for active reading, computers have recently become comparable to paper through replicating the latter's affordances. But even paper is not a panacea; it offers an inflexible document representation that supports some things well, such as embellishment, but supports others very poorly, like comparison and large scale annotation.

In response to this, I developed a prototype system, called LiquidText, to embody a flexible, high degree-of-freedom visual representation that seeks to alleviate some of the problems in paper and paper-like representations. To provide efficient control of this representation, LiquidText runs on a multi-finger touch and gesture based platform.

To guide the development of this system, I conducted a formative study of current active reading practice. I investigated knowledge workers' active reading habits, perceptions, and the problems they face with current reading media. I also inquired into what they would like in a future active reading environment. I used these results in conjunction with multiple design iterations and formative system evaluations to refine LiquidText for use in a summative study.

The summative study assessed, through a controlled, laboratory evaluation, LiquidText's impact on 1) the subjective experience of active reading, 2) the process of active reading, and 3) the outputs resulting from active reading. Generally, the study found a strong participant preference for LiquidText, and a focus on the creation of a summary of the original document as part of the reading process. On average, reading outputs were not significantly better or worse with LiquidText, but some conditions were observed that may help identify the subset of people for whom LiquidText will result in an improvement.

CHAPTER I
INTRODUCTION

Reading is not passive. Reading is highlighting, underlining, and annotating. It is showing relationships, taking notes in the margin, and elucidating the structure of a document. It's moving papers and lining up pages to compare just the right paragraphs with each other. It is jumping non-linearly between and within documents for comparison and search [O'Hara 1996; Liao, Guimbretiere et al. 2008]. Of course, some reading is passive—like perusing a novel. But the type I am interested in, known as *active reading*, constitutes a substantial portion of the time knowledge workers spend reading. Adler, et al. found that cross-referencing, in order to integrate information, alone constituted nearly 27% of reading activity for the wide variety of knowledge workers they studied [Adler, Gujar et al. 1998]. Reading to answer questions also constituted a significant fraction, at nearly 25%. As O'Hara discusses, both of these are very active reading tasks, often involving complex navigation, search, comparison, and annotation processes [O'Hara 1996].

Given this high level of interactive engagement that reading frequently requires, O'Hara and Sellen found that computers—at least in 1997—were not up to the task. Users experienced considerable difficulty with the spatial arrangement of their documents on the small, low resolution displays of the day. Annotation and efficient navigation of electronic documents were likewise problematic. Digital solutions simply could not provide the same level of flexibility and breadth of interaction affordances as their paper counterparts [O'Hara and Sellen 1997]. But in the time since their work was published, some of these concerns have been mitigated by advances in hardware and software. Ten years after O'Hara and Sellen's 1997 study, Morris et al. ran a similar evaluation to see if modern computers could better support active reading. They found that today's large, high resolution displays provide a comfortably large amount of room even to view multiple, full pages of text at once. Modern Tablet PCs allow virtual-ink annotation which similarly puts them roughly on par with paper. Indeed, such systems even support simple forms of bimanual interaction, where the user can operate a touch-sensitive scrolling area with the non-dominant hand [Morris, Brush et al. 2007]. In effect, these systems are becoming comparable to paper by replicating the affordances of paper. But this of course raises the question of how well paper itself supports the tasks of active reading.

Although paper is frequently seen as the gold standard for a variety of document-related tasks, I contend that, in many cases, paper may merely be better than the digital alternatives of the day, but not necessarily adequate to the actual task at hand. In order to consider paper's suitability to active reading, I refer to O'Hara's extensive taxonomy of the processes entailed by different reading tasks. While paper does provide numerous valuable affordances, it nonetheless appears lacking in several areas pertinent to active reading. One area that is of particular importance is the spatial arrangement of the text [O'Hara 1996; O'Hara and Sellen 1997]. Arrangement is critical, for example, when comparing different parts of a document. While paper may be easily arranged in space, the arrangement of content within a document remains fairly rigid. Particularly for a bound document, comparing passages on different pages can easily become laborious, requiring frequent flipping through the text, and consequently placing a greater strain on working memory. And although difficult when comparing two parts of a document, attempts to compare 3, 4, or 5 pieces of text quickly become infeasible without manually copying all the relevant content to another piece of paper. Unbound documents may make this easier in some cases, but this solution still scales poorly; as one seeks to compare text on 4 or 5 pages at once, the constraints of desk space alone begin to pose a problem. Moreover, arranging the papers to as to support comparison ruins their natural ordering, giving the reader a reassembly task before they can read the document linearly again.

The spatial rigidity of paper also makes it difficult to get an overview of a document. O’Hara and Sellen discuss the importance of this, and how, in the past, paper supported overviews better than computers [O'Hara and Sellen 1997]. But paper continues to offer little opportunity to step back and see the larger structure of the text—especially a lengthy one. Tables of contents offer some help, but they only refer to the text as originally created, and do not reflect reader-created notes and annotations that are a vital part of the active reading process [O'Hara 1996]. Notably, there have been some digital solutions to provide visual overviews of a document, either with outlines or overview-plus-detail displays [Greenberg and Gutwin 1996; Jakobsen and Hornbaek 2006]. But as I discuss in a subsequent section, these are generally useful but isolated systems, and do not exist as part of a larger active reading environment.

A related shortcoming lies in the navigation of a text—also known to be an important aspect of active reading [O'Hara 1996; O'Hara and Sellen 1997; Morris, Brush et al. 2007]. In some ways, paper provides good navigational affordances: its tangibility and amenability to elaborate bimanual interaction can make some navigation tasks quite efficient [O'Hara and Sellen 1997]. However, more complex forms of navigation, such as cross-referencing, pose more of a problem, as paper does not allow the user to create hyperlinks or otherwise create navigational relationships within the text.

Perhaps the most significant of paper’s limitations lies in annotation. In some ways, paper allows for flexible annotation; one can write small notes or comments anywhere on a page—even over text—without disturbing the text itself. And the annotations, being hand-written, can be naturally idiosyncratic and readily distinguished from the original content [O'Hara and Sellen 1997]. But for other types of annotations, paper is less adequate. By dividing text into pages at relatively arbitrary intervals, paper makes it difficult to create annotations that refer to large portions of the text, like a discussion that spans several pages. Perhaps more interestingly, it is likewise difficult to express a relationship between two pieces of text when they lay on different pages. It would not be hard to modify existing word processing software to support inter-page annotations, but it would still leave the problem of efficiently navigating to the multiple relevant areas of the text and viewing them simultaneously in order to create the annotation. Fundamentally, word processors and other modern text viewing software generally build upon a paper-like metaphor, and thereby inherit many of its shortcomings.

There are a variety of other issues as well. Paper places constraints on the type, and extent, of annotations which can be taken alongside, let alone inline with, the text. Computers, again trying to emulate the basic spatial properties of paper, often do little better. This is not to challenge the usefulness of paper, or some of the excellent affordances it does provide, such as being inexpensive and light. But as I discuss in more detail in the subsequent sections, paper is not a panacea—even for the things at which it traditionally bests computers. I propose instead that paper is rather inflexible; it assumes a particular way of presenting a document and while it allows one to embellish upon it, it provides little room to alter or restructure this presentation.

To summarize then, paper essentially provides a very stable, but rigid, form of representation. Generally, text remains in the same place relative to other text or annotations. The document is monolithic. In contrast to this rigidity, I propose to approach the representation of a document from a fundamentally different perspective: to offer a highly flexible, highly malleable document representation which the reader can dynamically tailor to suit the needs of their ongoing sense-making. Such a system, for example, should enable the user to rearrange a document so relevant pieces of text can be visible at once, but without destroying the linear flow of the document or removing the context of a given piece of text. It should facilitate seeing an overview of the document, but while providing needed detail. Likewise, annotation shouldn't be hindered by the scale or other spatial qualities of the document.

Thus, I have designed and implemented a prototype system that offers a fluid-like representation of a text where the user may restructure, re-visualize, and rearrange content to suit their needs. This system, called LiquidText, has served as a vehicle to enable me to evaluate and refine my conception of flexible document representations. LiquidText took shape as a document reader that incorporates a highly malleable visual representation, designed particularly to address some of the above limitations with viewing parallel sections of a text, reorganizing content from a text, and moving between parts of a text.

In a departure from the input devices traditionally associated with productivity work, I designed the LiquidText prototype around a modern multi-finger touch-screen, such as those described in [Han 2005] or [Wilson 2004]. As I describe in more detail in Chapter 4, this comes from the fact that the design for LiquidText offers a large number of degrees of freedom with which to manipulate the representation of a document. And while this representation is not based on a paper-like metaphor, it is founded on a *spatial* metaphor, where documents have geometric properties. Thus I believe, and further argue in Chapter 4, that multi-finger/multi-hand touch input is especially well suited to this application: it supports high dimensional input through the use of multiple fingers/hands in parallel, and also possesses a natural mapping to the spatial properties of the metaphor I employ. The former is especially important, as active reading is known to require parallel, and especially bimanual, interaction [Morris, Brush et al. 2007]. The latter opens the possibility for more natural, efficient control of LiquidText's complex, high degree of freedom representation [Norman 1988].

As part of the iterative design of the LiquidText system, I performed three studies: 1) a broad formative evaluation of an early LiquidText prototype, 2) a formative re-evaluation of LiquidText's gesture vocabulary, and 3) a summative evaluation comparing LiquidText to more traditional media. In the first of these studies, I investigated existing active reading practice, as well as performed a laboratory evaluation of the prototype system itself. My interest in obtaining new, empirical documentation of user behavior stemmed in part from the age of much of the existing reading literature. O'Hara's typology of reading goals, for example, is well over a decade old, and draws on even older sources, leading to concerns that the uptake of computer-mediated reading in the intervening years may have changed user behaviors and preferences [O'Hara 1996]. Secondly, the existing literature focused largely on current practice, rather than ideal practice. While understanding current active reading behavior is vitally important, I also sought to understand where readers experience difficulties and failures, and how active reading might be improved. Obtaining this data allowed me to better understand what functionality is most important for LiquidText, and to potentially help avoid some of the failures of other media.

While the first formative study offered many clear opportunities to refine the prototype, the solutions to some of the design challenges that emerged were less clear, and thus required many additional design iterations. The multitouch gesture vocabulary used by LiquidText was the foremost among these challenges, and I thus evaluated numerous small changes to the LiquidText gesture set. Fundamentally, the objective of this study was to ensure that participants could effectively learn and use LiquidText in the summative study, and thus provide a sound evaluation of the concept behind the system.

Finally, the culmination of this research was a summative evaluation of the system, where I ran a controlled study comparing active reading performed using LiquidText to the same tasks performed with traditional media. The value of this study was first in allowing me to identify LiquidText's impact on the reading process, and whether it could be a valuable complement or supplement to the tools traditionally used in active reading. More importantly though, it offered an opportunity to evaluate the concept of highly flexible document representations which underlies LiquidText, and thereby to make more general claims of how this concept could be applied in future research.

1.1 Research Questions and Thesis Statement

As I discussed above, underlying the design and development of LiquidText is a thesis about the nature of active reading. In this section, I explain this thesis, and discuss each of the research questions which I addressed in order to evaluate it (see Figure 1.1 for a summary of research questions and how I addressed them).

As discussed above, when doing active reading, people perform a variety of actions that depend on the spatial properties of the documents at hand. Some of these actions include viewing disparate areas of a text, commenting on large or non-contiguous portions of a text, acquiring an overview of a text, fluidly navigating a text, as well as skimming, searching, and the like. As I expounded upon, and as supported by my formative studies, some of these and other actions in active reading are cumbersome to perform. The rigidity of paper and paper-like document representations tends to make them inefficient and burdensome. Therefore, with these considerations in mind, the thesis which I will investigate throughout this dissertation is a follows:

> *Active readers require more spatial flexibility than is available in paper or present computer-mediated reading environments. Therefore, giving readers a computer mediated environment with a high degree-of-freedom visual, spatial representation, along with a comparably high-dimensional input method, will result in a subjective and objective improvement in the active reading process*

As the system I created, LiquidText, was intended to support the same tasks people already perform while active reading—such as navigation, annotation, obtaining overviews, etc.—I hypothesized that the type of improvement users would experience would be one of efficiency.I expected, however, that the exact cause of the efficiency increase would depend on the active reading task involved, and the media to which LiquidText was being compared. For example, comparing disparate parts of a document could be faster in LiquidText than with paper, because the former doesn't require time spent page flipping back and forth. But comparing disparate parts of a document may be faster in LiquidText than a *traditional computer application*, because LiquidText supports more input data to be supplied in parallel.

So in order to better understand and support the active reading process generally, and to evaluate my above thesis in particular, I investigated the following five research questions. Two of these are aimed at understanding the process of active reading as it stands today, the rest are to understand and assess LiquidText and its impact on reading.

1.1.1 RQ1: What are the circumstances surrounding active reading as it stands today, and what lower-level tasks do people perform as part of the active reading process?

Active reading is a broad category that encompasses a wide variety of reading activities, often with very different objectives and procedures [O'Hara 1996]. Nonetheless, certain lower-level tasks reappear across many reading activities, making them ubiquitous throughout the category of active reading—e.g., marginalia is used in both editing as well as studying. In order to support the active reading process, I began by identifying these constituent tasks. And in order to better support these tasks, I likewise sought to understand readers' motivations, goals, media, and other circumstances surrounding active reading generally.

1.1.2 RQ2: What difficulties do people tend to experience during active reading? Particularly, which aspects of active reading are most challenging?

As a wide variety of low-level, constituent tasks and processes appear throughout the category of active reading, understanding which of these tend to present the most difficulty enabled me to focus this intervention where its need is most acute.

In order to answer RQ1 and RQ2, I performed an empirical study of current active reading behavior across twenty-two knowledge workers from a variety of disciplines at a large manufacturing company. The study approached reading behavior from two perspectives: First, I sought to evaluate readers' situated, in-the-moment impressions of active reading using a diary study. Over the course of a week, participants answered a series of questions in a diary each time they completed an active reading activity. The questions elicited reactions while the memory of the reading was fresh in the participant's mind—before recency and primacy effects could interfere. Nonetheless, these diaries provided only a small cross-section of users' overall active reading experience. To get a broader picture, I also conducted interviews with individuals who had completed the diary study. Although subject to the challenges of recall, these provided a window into the ways in which active reading behavior changes between the different phases of people's work, and even the different phases of their lives (e.g., active reading in school versus professionally). It also provided an opportunity to probe in greater depth the motivations and difficulties associated with the constituent tasks involved in active reading. Together then, the two studies gave me a detailed cross-section of the tasks involved in active reading, as well as subjective motivations and reactions thereto.

The data gained through answering the first two research questions was used to revise the design of the initial LiquidText prototype. By contrast, the following three research questions focus on evaluating and investigating user reactions to the final LiquidText system.

1.1.3 RQ3: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the *subjective experience* of active reading?

Beyond just objective metrics, I sought to understand users' subjective impressions of the type of representation I use in LiquidText. This is essential to understanding factors such as motivation and affect, which could have longer term consequences for the overall value and utility of LiquidText and similar systems.

1.1.4 RQ4: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the *process* of active reading? To

what extent do users appropriate the affordances of the system, and to what extent do reader processes change in response?

In the formative evaluation of LiquidText, I performed a laboratory study in which participants completed an active reading task using the current prototype. The study showed that participants varied in the extent that they directly carried over their existing active reading habits and processes to their use of LiquidText. That is, some participants showed a greater affinity for reorienting their active reading behavior around the affordances of the system. Rather than a broad overview of the value of LiquidText, this evaluation revealed a detailed picture of the utility of LiquidText's specific functions for the given task. Thus, I performed a similar investigation in the summative evaluation to understand specifically which functions of LiquidText are used, as well as how those functions are used in readers' processes.

1.1.5 RQ5: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the results and outputs of active reading, vis-à-vis more conventional media?

While subjective opinions are important, active reading tasks frequently result in tangible outputs, such as summaries, critiques, and the like. And in order to assess the *outputs* of the active reading process, rather than *the process itself*, it was necessary to have a standard or benchmark. In this case, that standard was an output created while reading on more conventional media. Evaluating the differences between this and outputs produced while reading with LiquidText provided an objective measure of the impact of LiquidText on readers. To evaluate these outputs, I used a rubric tailored to the specific type of task participants performed.

Generally, to answer research questions 3, 4, and 5, I conducted a summative evaluation to compare participant reading behavior while using LiquidText versus reading behavior while using traditional media. The study began with a diary journaling task, akin to what was included in the formative study, where participants recorded their active reading behavior. The study then continued with a within-subjects, counterbalanced laboratory evaluation. Participants performd a control condition and an experimental condition, consisting of an active reading task performed with traditional media and with LiquidText, respectively. As part of the experimental condition, participants were taught how to use LiquidText, and were given several hours of practice time to encourage them to decide how best to appropriate the functions LiquidText offers. The study also included several questionnaires and interviews through which I assessed subjective participant reactions to LiquidText. And by designing the active reading tasks to include the writing of critical essays, I was able to compare participant outputs to assess their active reading performance across conditions. Finally, by observing participant habits and behaviors with traditional media compared to those with LiquidText, I was able to assess LiquidText's impact on active reading practice.

1.2 Contributions

The contributions of this work fall into two categories: first, my research helps to offer a comprehensive, contemporary understanding of the active reading process. Second, it offers a novel conceptual approach to supporting active reading, a concrete tool based on that approach, and an evaluation of the approach via the tool. Each of these provides generalizable knowledge that can benefit the research community. More specifically, the contributions include the following:

- The study of current reader practice offers a detailed picture of how people perform active reading activities. This helps to support the observations of earlier research, and also sheds added light on the novel problems that modern technology has introduced – e.g., the challenges of switching between digital and paper media. Additionally, prior research has generally focused more on the current state of active reading than on the problems people face, or on how people would like to resolve those problems. Identifying and elucidating problems, and exploring users' proposed resolutions, will help focus the research community on the most substantial issues that active readers face.

- Through evaluating LiquidText, this study provides a controlled case study of how people appropriate new types of affordances into established processes. The extent to which people adapted their active reading behavior in response to LiquidText helps provide an understanding of the malleability of processes such as active reading.
- Also through evaluating LiquidText, the research adds to the still nascent understanding of touch and gesture-based interfaces, particularly in the context of protracted knowledge work. As the community is still working to understand the implications of touch technology, and where best to apply it, this research helps to clarify the limits of touch interfaces as compared to paper and more traditional computer interfaces.
- Finally, the design ideas LiquidText is built upon offer the design community an alternate approach to document interaction. This comes partly from the specific interactions included in LiquidText, many of which could be applied in a variety of domains. But more generally, the underlying sort of flexible representation I use in LiquidText is quite atypical, and as such contributes to the general breadth of user interface design approaches. Also, as it lends itself especially to touch technology, it helps in the larger process of establishing interaction techniques for this relatively new type of platform.

1.3 Dissertation Overview

The remainder of this dissertation is organized as follows: In chapter two, I discuss background research and other work related to LiquidText. Chapter three discusses my background research into the active reading process, and answers the first two research questions. Chapter four explains the iterative design process used to develop LiquidText, and Chapter five discusses the resulting system in detail, including its contribution from a design perspective. Chapter six describes the final, summative study methodology, and chapter seven describes the findings of the summative study. Chapter eight is the conclusion and reviews the findings of the dissertation.

Research Question	**How I Addressed it**
1: What are the circumstances surrounding active reading as it stands today, and what lower-level tasks do people perform as part of the active reading process?	Diary study of active reading behaviors complemented with semi-structured interviews about the active reading process.
2: What difficulties do people tend to experience during active reading? Particularly, which aspects of active reading are most challenging?	Diary study of active reading behaviors complemented with semi-structured interviews about the active reading process.
3: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the *subjective experience* of active reading?	Interviews and questionnaires administered as part of a laboratory evaluation of the final LiquidText system.
4: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the *process* of active reading? To what extent do users appropriate the affordances of the system, and to what extent do reader processes change in response?	Laboratory evaluation of the final LiquidText system; particularly, comparisons of reading behavior while using LiquidText and while using more traditional systems.
5: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the results and outputs of active reading, vis-à-vis more conventional media	Laboratory evaluation of the final LiquidText system; particularly, comparisons of documents created by participants as a result of their active reading processes.

Figure 1.1. Strategy for addressing research questions.

CHAPTER II
RELATED WORK

In this chapter, I present the background of LiquidText by discussing prior investigations into related research areas, as well as specific projects that seek to provide similar functionality. This chapter is broken into three sections: First, I discuss background research on the topic of active reading generally, including people's motivations and processes for doing active reading. I then discuss specific research projects that have a similar goal to LiquidText, to support active reading, or some subset thereof. Finally, as LiquidText's design relies heavily on multitouch and gesture interaction, I provide a brief summary of related technologies in those areas.

2.1 Active Reading Background

The study of the process of reading has been an active research topic for well over a century. Early studies often investigated the perceptual component, carefully analyzing eye movements of readers, as well as various psychological and even philosophical questions about the process [Huey 1908]. Throughout the twentieth century, the study of reading became quite broad and quite deep as well. Researchers have looked carefully into the multitudes of facets of the reading process. Some, for example, have sought to understand the cognitive operations that underlie reading, and as such have developed various models of comprehension—e.g., [Perfetti, Rouet et al. 1999], [Samuels and Kamil 2002] or [Broek, Young et al. 1999]. There have also been extensive studies of reading focusing on some particular category of reading or reader. For example, Marshall et al. investigated legal research for the purpose of making a better e-book reader, [Marshall, Price et al. 2001]; Meyer explored a particular reading strategy in 9th graders [Meyer, Brandt et al. 1980]; and various researchers have explored issues of reading disabilities [Gough and Tunmer 1986; Kos 1991].

This proposal, however, is concerned with one category in particular—so-called active reading (although active reading does overlap with other categories as researchers often examine specific groups of readers from the active reading perspective, e.g., [Roast, Ritchie et al. 2002]). As discussed above, active reading includes a broad class of reading tasks characterized by a greater engagement with the text on the part of the reader. Some describe active reading as particularly involving critical thinking or learning [Schilit, Golovchinsky et al. 1998]. Others describe it as the process of creatively constructing one's interpretation of a literary work [Roast, Ritchie et al. 2002]. Adler describes reading generally as a continuum between the relatively passive and the active, but such activeness can manifest in a variety of ways, be they cognitive or physical [Adler and VanDoren 1972]. For the purposes of this work though, I particularly use "active reading" as referring to reading tasks that entail heightened physical engagement with media and tools, be they paper, highlighters, scissors, etc. While similar to that of Schilit and a subset of that of Adler, this notion of active reading involves not just cognitive activity, but places demands on the reader's environment—demands which can be more or less well met by the affordances that environment provides.

In offering a brief background of active reading, it is easy to get distracted by the incredible breadth of potentially related work, including cognitive psychology, linguistics, literary theory, studies of sense-making, and so forth. Thus, I will focus the discussion to the areas most relevant to my work: why people do active reading, what active reading consists of, and who performs active reading.

2.1.1 Why do active reading?

In order to understand active reading, and ultimately think about how to better support it, I will first consider why people do active reading—what goals they are seeking to accomplish. To investigate this issue, O'Hara developed a broad taxonomy of reading goals based on a survey of prior literature. From his descriptions of how people go about satisfying these goals, most would certainly fall into the category of active reading, including: *reading to learn; reading to search or answer a question; reading for research; reading to summarize; proofreading; reading to write and revise documents; reading for critical review; and reading to apply* [O'Hara 1996]. Murray performed a related study, and from that we might add the goal of *synthesis of new ideas based on a text* [Murray 2003].

While O'Hara and Murray derived lists of reading goals by consulting a variety of past studies, Adler et al. provided a unified, consistent empirical study of the topic, allowing for more quantitative results. By watching a diverse community of knowledge workers and categorizing their reading and writing behavior, they found 10 broad categories of reading, including some areas of potentially active reading not in the above taxonomies. Examples of those include: *skimming to decide whether to read the text in detail, reading for cross-referencing, reading to support discussion, or to support listening* [Adler, Gujar et al. 1998]. The highest frequency category was reading for cross-referencing, where information is being integrated from multiple documents, or multiple areas of the same document, at once. This constituted about 28% of time these knowledge workers spent reading and writing. Likewise, approximately 25% of the time involved reading performed to search or to answer questions, in which the reader is performing a goal-directed search that involves sampling a text to find answers or results.

These taxonomies are relevant to this research project because LiquidText is intended to support active reading as broadly as possible—that is, to support the common requirements that span these diverse reading goals. To this end, I next consider what actual tasks are entailed by these and other active reading goals.

2.1.2 What Constitutes Active Reading

In order to support active reading, we must understand more than just the larger goals people are trying to achieve. As described in Research Question 1, we need to develop a picture of the lower level tasks that people perform while pursuing those larger goals. But the individual low level tasks themselves are highly dependent on the details of the media in use—and the media in use in LiquidText (a flexible touch-controlled representation) is quite atypical. Thus we need to go past just what readers are doing, but learn why they are doing it; we need to know people's motivations and goals for performing lower level active reading tasks.

Naturally, the tasks entailed by active reading will vary somewhat depending on the larger goal at hand, the expertise and background of the reader, and other factors. The existing literature though, as well as my own studies, on active reading has shown a great deal of consistency among the tasks performed for different larger active reading goals.

To explore these tasks and motivations, I start once again with O'Hara's taxonomy, which broadly considers several types of reading and the processes involved in performing them. O'Hara describes a great breadth of practices, and one can see certain common themes, reflected throughout the literature, among the lower level tasks that comprise active reading. Two of the most prominent themes are navigation and annotation. In this section I will focus on these two categories of lower-level tasks, discuss readers' motivations for them, and point out some of the implications for design. I will more briefly discuss the theme of spatial arrangement in active reading, and content extraction and recording as well. Finally, I will give a cursory overview of several other lower level tasks involved in active reading as well.

The first theme I will consider is navigation, which is found across goals such as: reading to learn, to search, to answer a question, for research, for summarization, while writing, and for critical review [O'Hara 1996]. Naturally, all document reading includes the basic navigation of moving to a subsequent or prior page (or otherwise moving a predictable, incremental distance ahead or behind). But beyond this, active reading often includes a variety of important inter- and intra-document transitions; readers move between sections of a text often, sometimes to search or compare parts of a document, or to return to a place where they left off before a transition. Search tasks also entail non-serial reading, where readers may zig-zag through text looking for content [O'Hara 1996]. Similar phenomena were identified by Marshall and Bly in the context of magazine reading where, in addition to the expected page turning and flipping, readers would break the linear flow of their reading to look ahead for previews, and look behind for context [Marshall and Bly 2005]. O'Hara and Sellen discuss navigation as well in their observations of paper versus on-line reading. They noted how readers used bimanual interaction of paper interfaces to interleave and overlap navigation-related activities—such as maintaining a location with one hand, while flipping pages to search with the other [O'Hara and Sellen 1997]. Adler et al. similarly hint at the importance of navigation with their observation of the importance of cross-document content integration [Adler, Gujar et al. 1998].

Based on this and other related literature, there are numerous implications for the design of reading support systems. Some are explicit tasks that must be supported, and other are attributes of how those tasks should be supported. I list several of these here:

- Speed: navigation tasks must not interfere with the underlying reading task, and must be efficient [Lorch, Lorch et al. 1993; Sellen and Harper 1997].
- Breadth: to support tasks like skimming, large amounts of text must be traversable quickly [Muter, Latremouille et al. 1982; Marshall and Bly 2005].
- Metadata: in order to support rapid jumping through a text, readers need proper metadata cues suggesting what is where in the text [O'Hara and Sellen 1997; Marshall and Bly 2005].
- Dog-earing: readers must be able to quickly make short-term place-saving bookmarks to support navigation between prior and new locations [Askwall 1985].
- Cross-documents: readers must be able to navigate between multiple documents easily [Dillon, Richardson et al. 1989; Marshall, Price et al. 2001].

- Parallelism: navigation should not exclude other, simultaneous interaction with a text [O'Hara and Sellen 1997].

Another central component of active reading is annotation. I use annotation in this context to include any writing performed on the reading surface itself—typically, but not necessarily, in reference to some area of the text. This would include marginalia, highlighting, underlining, as well as various symbols and marks. Annotation, perhaps more than some other aspects of active reading, has garnered considerable interest in recent years, perhaps because of an early recognition that it must be supported in a variety of digital devices [Catlin, Bush et al. 1989], and perhaps because of its prominence and importance in knowledge work [Adler, Gujar et al. 1998; Wolfe 2000]. Since then, there have been numerous systems that support annotation in one context or another, for a variety of types of media.

My concern, though, is text documents. There, one of the earlier systematic, taxonomical studies of annotation was performed by Marshall, in order to design better annotation support for digital libraries [Marshall 1997]. Like [Nist and Kirby 1989], Marshall focused on students; she examined markings left in used textbooks across a variety of subjects, and was able to observe forms and infer functions for an array of different types of marks. The result was six overarching categories of function paired with the types of marks associated therewith. The functions included: *procedural signaling for future attention, place-marking and aiding memory, problem working, interpretation, tracing progress,* and *reflection on material circumstances of reading*. Building on this work, Renear at al. devised an extremely extensive taxonomy aimed at helping technology designers determine the types of annotation that systems ought to support [Renear, DeRose et al. 1999]. The kinds of annotation the authors discovered were organized along the dimension of function into six extremely broad categories (akin to, but broader than those developed by Marshall): *recording and scheduling reading, basic highlighting, commentary, classification, editing or coauthoring, and speech-acts*.

Although some research suggests readers' annotation habits can be translated quite smoothly from paper to tablet PCs [Marshall, Price et al. 1999], Qayyum developed an extensive taxonomy specifically on annotation in a digital environment. Surprisingly though, this was in the context of traditional desktop computers rather than in a tablet PC environment, which has been shown to provide an active reading experience comparable to paper [Morris, Brush et al. 2007]. Nonetheless, in developing his taxonomy, Qayyum ran a controlled study in which student participants used workstations with Adobe Acrobat to annotate school-related documents, either for their private or collaborative use [Qayyum 2008]. This is akin to a similar public/private comparison performed in [Marshall and Brush 2004]. The results are an in-depth taxonomy of the various markings along some exposition of the reasons for those markings. While not as theoretically rigorous as Renear et al., this taxonomy provides considerable insights and comparisons between traditional computer interfaces and paper. As one might expect from earlier work [O'Hara and Sellen 1997], PC workstation interfaces were considered to have better tools and better support archival and searching; however the reading and thinking process seemed to suffer.

As many of the studies in annotation are intended to promote better system design, they offer truly copious amounts of design implications. For the sake of space, I will briefly point out several of the most prominent themes:

- Idiosyncrasy: People use markings in complex, emergent ways that change with time or situation; as such, the annotation process must support rich, dynamic ways to mark documents [O'Hara and Sellen 1997; Renear, DeRose et al. 1999; Marshall, Price et al. 2001].

- Seamlessness: Especially when switching between different types of annotation, selecting and controlling functions must be highly efficient and not distract from the underlying reading task [O'Hara 1996; Marshall 1997; O'Hara and Sellen 1997].
- Separation: Annotations must be readily separable, temporarily or permanently, from the original content [O'Hara and Sellen 1997; Renear, DeRose et al. 1999].
- Retrieval: Effective ways to find and navigate among annotations are critical for certain goals, such as identifying text to be read [Renear, DeRose et al. 1999].

Navigation and annotation are perhaps the most prominent components of active reading, but other types of tasks are naturally involved as well. Somewhat related to navigation, spatial layout is also known to be an important part of this process [O'Hara and Sellen 1997]. In particular, layout affordances are especially crucial during certain types of tasks, such as: *gaining an overall sense of a document's structure, cross-referencing, and interleaving reading and writing* [O'Hara and Sellen 1997]. Without presenting a detailed analysis of the full breadth of tasks involved in spatial layout, I will briefly describe several implications for design that the literature suggests:

- Parallelism: Systems that support active reading must readily enable parallelism, both in viewing multiple documents and in viewing multiple parts of the same document at once [Lorch, Lorch et al. 1993; O'Hara 1996; Morris, Brush et al. 2007; Golovchinsky 2008].
- Linearity: While viewing sections of a text in parallel is valuable, it should not come at the cost of disrupting the underlying linearity of the text [O'Hara 1996].
- Speed: Spatial arrangement is a dynamic process, with content moving in and out of the reader's focus of attention; it must thus be efficient to rearrange the layout [O'Hara and Sellen 1997].

- Awareness: Spatial layout should support a sense of orientation within a document as well as in the larger space of documents the active reading task entails [O'Hara 1996; O'Hara and Sellen 1997].

Annotation, as I have defined it, involves marking on the reading surface; but another important part of active reading is writing on a *secondary surface* of some sort (e.g., note taking, outlining, etc.). Some reasons for this may include summarization of an original text, or integration of content from multiple sources [O'Hara 1996]. Readers also may simply prefer to take more consolidated notes in a separate document or computer application [Marshall, Price et al. 2001; Marshall and Bly 2005; Golovchinsky 2008], rather than in the original document. An empirical study of student reading behavior described in [O'Hara, Smith et al. 1998] looked particularly into one aspect of writing-while-reading: extraction and recording of information while reading. This would include tasks such as note taking and outlining. They found the students recorded four broad categories of material; this led the experimenters to identify a variety of specific motivations for recording content on secondary surfaces. Among these were: *facilitating concentration and memory, supporting later review, condensing content, using in later compositions, as a means of indexing into the source text,* and *to act as a memory cue for related information* [O'Hara, Smith et al. 1998; Marshall, Price et al. 1999]. We can add to that list: *integrating together information from multiple sources,* and *reorganizing the original text's content* from [O'Hara 1996]. Another study of information extraction, in the context of periodicals in particular, identified several additional motivations, including: *reminders for action*, and *sharing with others* [Marshall and Bly 2005]. Although there is not as copious a literature base discussing information extraction as annotation, the prior research does offer some implications for design, including:

- Traceability: it must be easy to determine the source or context of extracted text [O'Hara, Smith et al. 1998; Marshall and Bly 2005; Golovchinsky 2008].

- Review: readers must be able to readily review, manage and organize the content they have extracted [O'Hara, Smith et al. 1998; Marshall, Price et al. 2001].
- Efficiency: Extracting information should be fast and easy; it should be tightly integrated into, and not interfere with, the reading process [O'Hara 1996; Marshall and Bly 2005; Golovchinsky 2008].

Of course, more than just annotation, navigation, layout and extraction may be involved in active reading. For example, when the active reading process involves multiple texts, the reader can be faced with tasks such as organizing the various texts involved in order to facilitate retrieval and awareness [Marshall, Price et al. 2001]. If those documents, or some of the content within them, are meant to be used collaboratively, it can entail other tasks as well [O'Hara 1996; Qayyum 2008], especially in the context of digital libraries [Marshall 1997]. Generally, there are a variety of lower level tasks involved in active reading, and the purpose of this section has been to discuss the space of these tasks, point out some of the motivations behind them, and list some of the resulting implications for design. So having discussed why people do active reading, and what active reading consists of, I will next consider who does active reading.

2.1.3 Who Does Active Reading

Although active reading is principally associated with focused knowledge work, it is actually found quite broadly and in many contexts [Schilit, Golovchinsky et al. 1998]. Naturally, though, one of the most prominent groups would be professional, career knowledge workers. Adler observed active reading behavior across a spectrum of such workers, including lawyers, doctors, business people, and various others [Adler, Gujar et al. 1998]. Economists likewise fit this category and their work has been the subject of considerable analysis [Sellen and Harper 1997].Our understanding of active reading has also come from as administrators and research scientists, who are known to engage in these types of activities [O'Hara and Sellen 1997]. It is important to note, however, that even as many knowledge workers do active reading tasks, the goals behind those tasks vary widely, and seem to vary more by individual than by field of work [Adler, Gujar et al. 1998].

Students are also known to engage frequently in active reading tasks. The digital libraries community, perhaps because of students' considerable use of libraries for school work, has focused attention on studying their reading behavior. A considerable body of literature thus examines them, including [Marshall and Brush 2004], which looked at graduate HCI students from various backgrounds, and [Marshall, Price et al. 2001] which specifically observed law students. A study of a variety of students is found in [Marshall 1997], which analyzes annotation in textbooks across a range of disciplines.

Active reading is also found in more personal reading tasks. In studying people's use of traditional periodicals, Marshall and Bly found that people performed a wide variety of complex interactions with their reading media, including extracting and attempting to re-contextualize content, sharing content, and various non-serial navigation tasks [Marshall and Bly 2005; Marshall and Bly 2005].

A key message then, of past research, is that active reading is encountered not only in rare, specialized domains. Rather, it is a part of an extremely broad range of knowledge work, even outside of the professional sphere.

2.2 Related Projects

With the prevalence of active reading across so many disciplines and groups, many projects, both research-oriented and commercial, have endeavored to offer better support than traditional computer-based approaches. That traditional computer systems fall short was demonstrated in [O'Hara and Sellen 1997], which showed that paper possesses numerous affordances that computers lacked, at least in 1997. Since then,

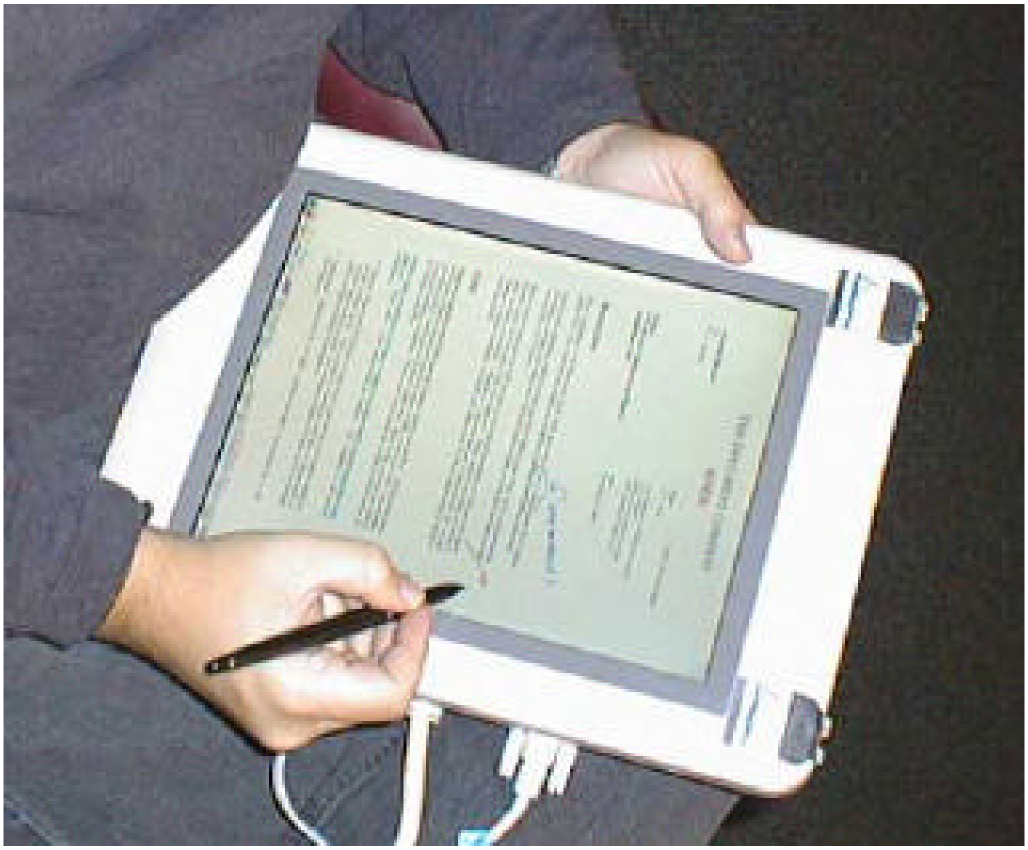


Figure 2.1. XLibris system in use.

commercial technology has improved. Modern tablet computers include many of the affordances of paper; and their software supports freeform inking that makes for an active reading experience comparable to that of paper itself [Morris, Brush et al. 2007]. But along with this commercial improvement, the research community has sought to go even further toward matching, or even advancing past, paper. In order to better explain what LiquidText is, and how it contributes to this field, I will consider previous research projects with similar goals, and point out how they relate to LiquidText.

2.2.1 Active Reading Systems

To begin the discussion of past work, I will consider the projects more explicitly aimed at supporting active reading, as opposed to those that only have incidental feature overlap with LiquidText.

Among these, the most prominent and complete system in the literature is likely XLibris [Schilit, Golovchinsky et al. 1998] (Figure 2.1). XLibris is an early tablet computer-based system that uses an electronic pen to interact with documents. It was intended to out-perform paper for active reading tasks, by starting out with an explicitly, extremely paper-like metaphor, and combining that with a variety of digital functionality. Generally, XLibris supports highly unstructured input, focusing especially on free-form ink annotation. From this, the system can make simple inferences about the importance of different regions of the text and, in turn, suggest other possibly relevant documents. XLibris will also aggregate annotated text into a Reader's Notebook, where the user may browse all of their annotated content in one place. An evaluation of the system with a reading group suggested that users thought of, and used, the system very much like paper [Marshall, Price et al. 1999], indicating that XLibris was generally successful in its use of a paper-like metaphor.

Others have used XLibris as a starting point for exploring additional active reading support functions. Price et al. incorporated manual linking into XLibris, allowing the user to make *ad-hoc* markings that act as anchors, and which can be joined to one another, allowing disparate parts of a document to be associated[Price, Golovchinsky et al. 1998]. A more comprehensive redesign of XLibris came while trying to make the system more useful for legal research [Marshall, Price et al. 2001]. After carefully observing legal students' active reading habits, the researchers added a variety of new features for backtracking and navigation between document views, and for supporting keyword searchers and navigation to web pages. To the "Reader's Notebook" of annotated content, the redesign included ways to further annotate content while it was being viewed within the Reader's Notebook, and more customizable ways to organize the extracted content as well.

Although XLibris, like LiquidText, includes a rich complement of functionality to support active reading through annotation, navigation, and the like, its differences from

LiquidText are quite fundamental. XLibris was designed explicitly to provide a paper-like metaphor, as can also be seen readily from the system's design. But as the authors observed, even fully replicating the affordances of paper is challenging. Users, for example, regretted the lack of flexibility in not being able to fold or otherwise spatially alter pages to construct different, parallel representations of content. They also found that the Reader's Notebook was too cramped, diminishing its usefulness [Marshall, Price et al. 1999]. Generally, XLibris suffered from the lack of flexibility available in its visual representations. Perhaps epitomizing this was the observation on the part of some users who felt that while XLibris was very comfortable to use, its advantages over paper were unclear. One user explained it thus:

> "I could have more from this device, because it was too much like plain, ordinary paper. And there must be a high powered computer behind it. But I wasn't really taking advantage of the power."

By contrast, LiquidText does not attempt to be paper-like. Rather than adding digital functions to a paper-like metaphor, LiquidText seeks to provide a different, more fluid metaphor based on representational flexibility. As a necessary result of this, LiquidText is also designed around a very different form of input than XLibris and other E-book systems. XLibris' pen input, while natural, misses the bimanual interaction that is an important part of reading a paper text [O'Hara and Sellen 1997]. While some e-book systems do support simple forms of bimanual interaction, such as scrolling [Morris, Brush et al. 2007], the use of a multi-touch display allows LiquidText to offer the potential for richer multi-handed input.

XLibris sought to mesh a paper-like representation with the power of modern computing by using a paper-like metaphor on a small, light computer system. Others, however, have gone even further to maintain paper's affordances, and actually kept paper as the input medium. Among these projects are the DigitalDesk [Wellner 1993], which allowed the mixing of ink and computer-based projections all on regular sheets of paper. Also included would be EnhancedDesk [Koike, Sato et al. 2000] which could augment a

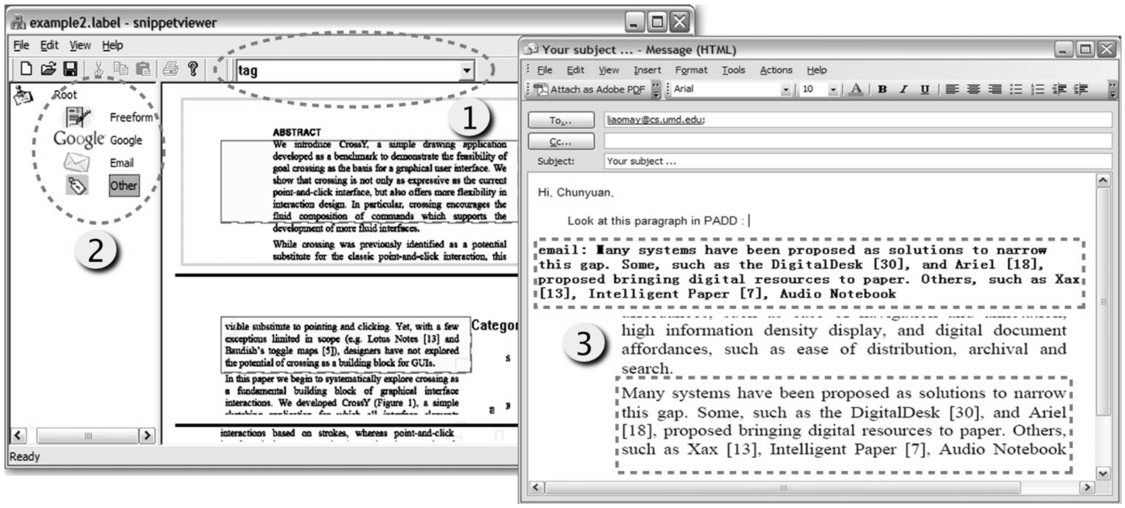


Figure 2.2. PapierCraft's software interface for managing snippets.

textbook with digital content, as well as Video Mosaic [Mackay and Pagani 1994] which supported computer-based video production while integrating with paper storyboarding. However, the systems most closely associated with active reading would be PADD [Guimbretiere 2003] and its successor PapierCraft [Liao, Guimbretiere et al. 2008] (Figure 2.2). This project uses the Anoto system to record all of the markings made on a sheet of paper, and offers a rich array of pen-gestures to enable actions to be denoted on paper, and automatically performed on the digital copies, of a document. In this way, the system also captures free-form annotations which can be made on paper and transferred to digital [Liao, Guimbretiere et al. 2008]. PapierCraft includes a variety of other features, but essentially approaches the problem of active reading support in a similar way to XLibris. PapierCraft explicitly seeks to enable users to transfer their paper practices to this new system, and to provide an enhanced paper-like experience. LiquidText, although trying to support similar activities, is of course trying to do so through a different approach.

2.2.2 Other Systems

In contrast to systems like LiquidText, XLibris and PapierCraft, a variety of research projects have addressed issues of document annotation, navigation, re-representation, and the like, without seeking to support the whole of active reading, or even active reading *per se* at all. In order to better situate my research in the larger field of related work, I will note several of these systems and describe their relationships to active reading.

Navigation is one of the central categories of active reading tasks, as described above, and hypermedia systems have been exploring navigation extensively. Specifically, hypermedia supports the presence of elaborate networks of links and other relationships between pieces of content. The MetaLinks "hyperbook" system takes this approach, and seeks especially to support active reading. It allows authors to create content containing complex sets of various types of links. The unique structuring of the links enables the reader to view the document with considerable navigational flexibility, traversing the text at different levels of generality, or jumping to associated content [Murray 2003]. As navigating to associated content in hypermedia systems can impede semantic coherence [Foltz 1996], MetaLinks also enables authors to create bits of text to help smooth those transitions. While hypermedia systems such as MetaLinks have considerable potential, they employ a very different approach than LiquidText. One important difference is that systems such as these require content that has been especially created for them—so they include complete content creation/consumption systems. LiquidText, by contrast, is designed to help readers of even traditional content without any author-added metadata.

Another approach to supporting navigation is found in the Fluid Documents project [Bouvin, Zellweger et al. 2002]. Generally, Fluid Documents "explores how electronic documents can provide more details on a page by fluidly adjusting typography on demand [Zellweger, Regli et al. 2000]." To this end, the project demonstrates several approaches to displaying bits of content related to a particular part of a document. For example, a definition of a word in a document might be selectively displayed by shifting the rest of the document down to make room, by translucently overlaying the definition on top of the document text, by showing it in the margin, etc. This can support navigation by providing an in-place means of accessing various forms of metadata associated with document content, such as links to, or previews of, related content. Fluid Documents describes various applications of these techniques, as well as an architecture whereby different pieces of content can even negotiate for space [Chang, Mackinlay et al. 1998]. Although Fluid Documents is not designed as a general active reading support system, it does offer interesting ideas displaying metadata and document links, and as such may be of value to LiquidText as well.

As discussed above, representational flexibility is an important part of active reading—one LiquidText focuses on. However, many other systems have explored novel approaches to the visual representation of documents or text. One area that has especially been the target of this research is software development—helping programmers better visualize code. Systems such as SeeSoft [Eick, Steffen et al. 1992] and Tarantula [Jones, Harrold et al. 2001] provide hierarchical zoomed-out views and color coding to provide a usable overview of very large code documents. As is often the case with providing overviews, others have explored integrating them with detail views using fisheye representations of source code [Jakobsen and Hornbaek 2006]. Many other systems such as [Reiss 2006] operate on higher-level semantic units and display elaborate hierarchies of procedure calls or other relationships.

Outside of source code, elaborate text document visualizations are perhaps less common, but still explored in a range of projects. One example of this is VIKI, which uses “spatial hypertext” to support elaborate hierarchical organizations which a user may express through spatial relationships [Shipman, Marshall et al. 1999]. As part of this work, the authors explored different ways to represent large numbers of content-representing nodes so as to preserve spatial context, yet while ensuring certain nodes remain fully readable. To this end, they explored several fisheye-type approaches to visualizing these nodes and their relationships. Although not a complete active reading support system, this type of work was used by the authors for visualizing relationships between documents, and could be used to visualize relationships between pieces of documents as well—all of which are important parts of reading.

More closely related to active reading however, [Hornbaek and Frokjaer 2003] explores several visualization schemes for text documents, and their impacts on users. They look at traditional text rendering, and two types of distortion similar to those used in code visualization. The first distortion is an overview + detail view where, parallel to the main text, is a scaled down, or otherwise condensed, view showing an overview of the document, as in [Ginsburg, Marks et al. 1996] or [Egan, Remde et al. 1989]. The second distortion is a fisheye view of text, where some of the content is selectively scaled down to provide an in-place overview, such as in [Robertson and Mackinlay 1993] or [Holmquist 1997], which considers the use of semantic zooming as well as geometric zooming. This is especially relevant to my work as LiquidText makes use of overview + detail views, as well as a form of fisheye visualization. In their study, Hornbaek and Frokjaer found that along a variety of objective and subjective metrics, either the overview + detail or the fisheye view usually resulted in an improvement over a traditional, linear rendering. However it was less clear whether overview + detail or fisheye was superior [Hornbaek and Frokjaer 2003]. In LiquidText I offer a form of both views, intending to give the reader the flexibility to use the most appropriate view for the task at hand.

Document visualization has been explored in other areas as well, such as supporting awareness of a document’s history. Projects like Edit Wear And Read Wear, for example, indicate to the user the areas of a text that have been read or edited the most over time [Hill, Hollan et al. 1992]. This can be a powerful cue in aiding users in finding the most salient areas of a text; however, to keep the project focused, LiquidText is not intended to address collaborative scenarios at this point. But in the longer term, cues like these could play an important role in the future development of LiquidText.

Another aspect of active reading that has received considerable attention is annotation. One type of annotation is inline, associated content, and one problem is how to position this content in a document—e.g., where should reader comments be placed. To address this, part of the Fluid Documents project, described above, included “fluid annotations” for web pages. This enables page content to be associated with user-selected text, pictures and links [Bouvin, Zellweger et al. 2002]. As with other parts of Fluid Documents, these annotations are normally hidden, but selecting the pieces of content with which they are associated will cause the annotation content to expand inline to full size, pushing down other lines of text as needed.

Fluid Documents does not, however, address other types of annotation, such as those more deeply integrated into the layout and structure of the document, such as underlining, or circling of text. These annotations are necessarily overlaid on document content, and so would not require document content to move out of the way to display them. However, their reliance on the spatial properties of the document also makes them unstable in the case of a text reflow. To address this problem [Golovchinsky and Denoue 2002] describes a system of heuristics for intelligently adjusting a variety of such page markings so they correspond to the same text even after a reflow. Although this idea, as well as “fluid annotation,” above, are not presently necessary given LiquidText’s current functionality, I expect to consider them as I refine LiquidText’s design, as discussed below.

2.3 Touch and Gesture Interfaces

As described above, LiquidText uses multi-finger touch screens as its primary input device, as this affords high dimensional, spatially mapped input. While I believe LiquidText to be the first touch and gesture-based system specifically designed to support text manipulation in general, or active reading in particular, this type of input has been applied and investigated in a variety of other domains. In this section, I will provide a brief overview of multi-touch and some of the systems that employ it.

2.3.1 Sensing

While touch sensitive screens have been used in various portable and kiosk applications for many years, these systems have generally been able to sense only a single touch-point. With rare exceptions (e.g., [Loviscach 2007]), placing multiple fingers on these sensors would simply corrupt the input. More recently, though, a variety of approaches have been developed to sense two or more fingers touching a screen at once.

Many of the most popular approaches to multi-touch sensing are optical—that is, they seek to acquire an image of the surface of the display, which is then processed using computer vision software to identify fingers or other objects. Perhaps the most popular of these is frustrated total internal reflection (FTIR). These systems involve rear-projection displays with precisely arranged lighting and cameras to identify fingers or other objects contacting the screen [Han 2005]. An approach developed by Wilson is similar in its use of rear projection, but has the potential to detect objects in proximity to, but that do not touch, the display [Wilson 2005]. More recently, a new class of optical touch screens are exploring the use of LCD, rather than rear-projected, displays. As a result, some of these systems can be built extremely thin [Izadi, Hodges et al. 2007], or at least dispense with the added cost and heat of the projector [Motamedi 2008].

A fundamentally different approach to touch sensing that naturally lends itself to thinner, lighter systems is capacitive. While capacitive multi-touch screens are relatively new, single-finger capacitive touch screens as well as multi-finger, opaque touch sensors have been available for many years [Lee, William et al. 1985; Evans 1989]. One of the earlier multi-touch displays was found in DiamondTouch, which used front-projection to form the image on an essentially opaque capacitive multi-touch sensor [Dietz and Leigh 2001]. More recently, transparent capacitive multi-touch sensors have been developed commercially[1] and can be used with LCD displays. However, these thin, light multi-touch screens are presently very difficult to construct at sizes above twenty inches[2], making them more appropriate for smaller, portable devices. Another shortcoming of capacitive sensors is that they usually can only detect skin, and not fingernails or other objects that may touch the screen. Worse yet, they generally cannot provide the complete shape/silhouette of the object touching the screen, making it difficult to support whole-hand interaction. In spite of these shortcomings I have so far been using capacitive touch sensors with LiquidText due to their small form factor which allows them to be easily moved and angled by the user.

2.3.2 Multi-touch Systems

Although I am aware of no other multitouch systems designed especially to support active reading, there have been various other forays into knowledge work using multitouch interfaces. Perhaps because multitouch lends itself to the direct control of concrete visual objects, some past research has explored its use in design. Gingold et al explore the use of multitouch for controlling texture placement on 3D surfaces [Gingold, Davidson et al. 2006], and [Gallardo, Julia et al. 2008] demonstrates a touch/tangible

---

[1] Commercial systems are available from N-trig or 3M, at www.n-trig.com or www.mmm.com.

[2] As explained in private correspondence with an N-trig executive.

variant of the Logo programming environment. Projects such as [Wu and Balakrishnan 2003] and [Rick, Harris et al. 2009], for example, have looked at multitouch design support from a collaborative perspective, although for very different user groups. The theme of collaboration support is also found in numerous other projects, such as [Wang and Maurer 2008], which supports project planning, [Jiang, Wigdor et al. 2008], which supports collaborative scientific work, and [Apted, Sydney et al.], which helps people to share photos. Closer to the vain of LiquidText, some work is even being done in helping groups of people to do document editing in parallel [Ringel, Ryall et al. 2004; Masoodian, McKoy et al. 2007], or group-based school assignments [Piper and Hollan 2009]. However, none of these systems are intended to support active reading; rather, the first two act as intermediaries for collaborative document use, and the third is for performing visual, diagram-centric classroom activities.

While multitouch may be valuable for collaborative work because multiple people can readily use such systems at once, the considerable flexibility is affords has led to various personal expression applications as well. Music synthesis and control, for example, has made use of this technology through several projects, such as [Davidson and Han 2006; Bredies, Mann et al. 2008; Lemur 2009]. Similarly, some projects attempt to explore visual aesthetics as well, such as [Schiphorst, Motamedi et al. 2007; Vandoren, Claesen et al. 2009]

Although there are numerous multitouch projects outside the domains of design, collaboration and the arts, the technology is sufficiently nascent that we are only beginning to figure out its applications. But an important resource for creating applications is a rich array of interaction techniques appropriate for the platform. Here, we find, for example, considerable research has been devoted to mitigating the disadvantages of going from the traditional mouse/keyboard to multitouch, such as precision pointer input [Benko, Wilson et al. 2006; Esenther and Ryall 2006], and text entry [Hinrichs, Schmidt et al.; Hirche, Bomark et al.; Hinrichs, Hancock et al. 2007].

Others have sought to better understand the proper gestures to use for different purposes, and some have even performed systematic investigations of users' gesture preferences [Grossman, Wigdor et al. 2005; Epps, Lichman et al. 2006; Davidson and Han 2008; Wobbrock, Morris et al. 2009].

2.3.3 Multitouch Theory

Although multi-point input has been explored for many years [Bier, Stone et al. 1994], only recently have true multi-finger touch screens become widely available. As the previous section shows, this technology can be applied in a variety of ways for many different purposes. In this section I will break down the basic uses of multitouch and point to some of the systems that employ them.

One important use of multitouch is to achieve parallelism. In this context, parallelism refers to support for the user's performance of multiple actions at the same time [Balakrishnan and Hinckley 2000]. Clearly this can be beneficial, as any overlap in task performance can result in a reduced total completion times, compared to serial performance. Though, of course, in some cases it is possible for the cognitive burden associated with simultaneous action to partly or completely outweigh the performance gain [Kabbash, Buxton et al. 1994]. Within the category of parallelism, the various parallel inputs (such as a user's two hands, or several fingers) may be working *cooperatively* or *independently*. Cooperative performance would have the inputs working together to accomplish a task—such as opening a jar or untangling a knot. Independent performance would have the inputs performing separate tasks entirely—such as controlling different sliders of an equalizer. Here, I consider these two categories in more detail.

Conceptually, parallel-*independent* interactions with multitouch are perhaps the most straightforward. Interfaces can behave much as they would with a mouse, except that more than one widget or object can be controlled at once. We see this in systems like

[Davidson and Han 2006], which includes support for each of a user's hands independently controlling knobs and sliders on an audio synthesis system, and in various collaborative systems [Wu and Balakrishnan 2003; Piper and Hollan 2009].

In contrast to parallel-independent interactions, parallel-*dependent* interactions involve the different inputs (again, such as a user's two hands) cooperating to perform a task. However there are different ways to cooperate, which are known as *symmetric* and *asymmetric* interaction [Balakrishnan and Hinckley 2000]. In this case, the notion of symmetry refers to the role that one input plays relative to the other. In asymmetric interaction, the inputs (usually hands) play complementary but disparate roles; so one input's role must be performed in order for the other input to perform its role. For example, when opening a jar, the hand grasping the lid cannot perform its role of rotation unless the non-dominant hand holds the jar in place. By contrast, symmetric interaction involves the inputs performing similar but independent actions to accomplish the same task, such as specifying opposite corners of a rectangle [Casalta, Guiard et al. 1999].

Part of the interest in categorizing parallel interaction according to symmetry is because asymmetric interaction can be readily modeled through what is known as the kinematic chain theory (KC). KC provides a way to understand the roles of the different components of a kinematic system such as an arm. It shows, for example, that distal components operate in the reference frame established by proximal components, that increasingly distal components operate on increasingly small scales, and that distal components mobilize after proximal components [Guiard 1987]. The value of this model is that some bimanual interactions can be treated analogously—where the non-dominant arm acts as a proximal component in a KC model, and the dominant arm acts as a distal component. Thus, in the case of these cooperative asymmetric tasks, (assuming right-hand dominance) the left hand tends to set the reference frame for the right hand, the left tends to mobilize before the right hand, and the left tends to operate on a larger scale than the right [Casalta, Guiard et al. 1999]. This model gives insight into a common mode of parallel interaction, and as such can help inform designs of multitouch systems. We see multitouch interactions like these discussed in [Benko, Wilson et al. 2006] for use with precision selection, and in [Wu and Balakrishnan 2003] for use with rotation.

Symmetric interaction, again, involves the different inputs independently performing the same type of action to accomplish a common purpose. One example of this can be the common scaling gesture as seen in [Wu and Balakrishnan 2003], or in the control of texture placement nodes in [Gingold, Davidson et al. 2006]. Note that in spite of humans' considerable experience performing asymmetric interactions in our real-world activities, the choice of whether to design an interaction to be symmetric or asymmetric depends very much on the details of the task—as symmetric can sometimes be faster [Casalta, Guiard et al. 1999].

Of course, in addition to all of the above parallel interactions, another category of multitouch interaction is not about taking advantage of parallel inputs, but about greater expressiveness in each input. For example, referring again to [Wu and Balakrishnan 2003], we see the side of a hand has different semantics than the tip of a finger. Even though both represent just one input, the side of the hand can be used to "sweep" objects shown on the display, in contrast to the use of the fingertip. Similarly, we see a variety of different meanings associated with different hand shapes in [Epps, Lichman et al. 2006], and even different finger configurations in [Benko, Wilson et al. 2006]. Although not all multitouch hardware can accurately detect complex hand configurations yet, these techniques do help to further increase the bandwidth of multitouch input.

CHAPTER III
UNDERSTANDING ACTIVE READING

While the prior research provides extensive detail on the active reading process, there are several significant gaps in this literature. The first is age. Although there has been some recent work on understanding active reading (e.g., [Morris, Brush et al. 2007; Golovchinsky 2008]), much of the core research in the field was performed well over a decade ago (e.g., [O'Hara 1996; Adler, Gujar et al. 1998; Marshall, Price et al. 1999], etc.). However, much has changed since that time; computing has taken on a broader role in content consumption, and new computing form factors have become common. As a result, our understanding of active reading may be partially obsolete.

The second gap is in our understanding of the problems people face in active reading. Over the past two decades, there have been various empirical explorations into why computers tend to underperform paper as an active reading medium, and how computers may be better adapted to match paper's performance (e.g., [O'Hara and Sellen 1997; Morris, Brush et al. 2007]). Remarkably though, there has been little investigation into the general problems that readers face irrespective of medium—particularly, the problems that readers face apart from PC's, such as with paper. The general literature on the reading process does allow one to construct a possible theoretical argument about what paper may lack, but little empirical investigation has been performed. Thus, we are left in a position of uncertainly when trying to create active reading technologies that surpass the status quo.

I sought to address these two gaps, which correspond to RQ1 and RQ2 respectively, by conducting a formative study in which I learned from a variety of participants about their reading habits, preferences, challenges, and ideas. As a step toward answering the remaining research questions, I also conducted a formative evaluation of the then-current LiquidText prototype intended to inform the redesign of the system. Given the considerable methodological overlap between these two studies, I integrated them into a single formative study. In this chapter, I discuss this study and its results as they pertain to RQ1 and RQ2, as well as to participants' general thoughts on the future of active reading. I will discuss the formative evaluation of LiquidText in the subsequent chapter.

3.1 Methodology

Broadly, I gathered data for this study in three ways: 1) I conducted a diary study in which participants recorded their reading behaviors, challenges, and preferences; 2) I conducted interviews, pursuing the same topics in a less structured way; and 3) I conducted two design workshops in which participants envisioned what they would want from a future active reading environment. These three components help provide different perspectives on readers' behaviors, needs and ideas. The diary study let me gather in-situ, naturalistic data about readers' active reading situation over an extended period of time, telling me about their context, choices of tools, and the like—which would be difficult to obtain using laboratory methods. The interviews offered a way to get retrospective feedback from readers, as they generalized over the data they collected though their diaries. And the workshops gave readers a chance to think collaboratively, integrating their personal experiences and ideas with those of others.

While active reading possesses different connotations in different contexts, I chose a broad lens in order to identify active reading across a wide range of scenarios. This broad definition was to focus on reading activities that involve more interaction with the reading media than simple sequential advancement through the text (like page turning). This includes, for example, searching a text, comparing texts, annotating, bookmarking, etc., and is consistent with the usage of the term *active reading* in related literature [Schilit, Golovchinsky et al. 1998; Morris, Brush et al. 2007; Liao, Guimbretiere et al. 2008]. Notably though, that the boundary between active reading and the many other tasks that accompany reading is not always clear. For example, active reading often occurs intertwined in a larger workflow such as composing a report, studying for a test, or attending a meeting. To maintain a clear distinction, I define active reading *per se* as the part of the workflow involving the actual reading medium itself.

3.1.1 Participants

With notable exceptions, many studies of active reading focus on fairly narrow subject pools (e.g., students). So to investigate the subject generally, I recruited from a large, vertically integrated, design and manufacturing firm—which offered a broad pool, ranging from woodworking engineers to researchers. I spoke with participants before enrollment to screen out people who were not frequent active readers, and offered participants a $5 lunch voucher for their time.

Of the 24 people who initially volunteered, 20 completed the study, 13 of whom were women. The jobs of these 20 were as follows (note that the students were also interns):

Manager	P1
Design student	P2, P3, P5
Design researcher	P4, P6
IP Lawyer	P7, P9
Math/Computing student	P11
Mgmt info systems student	P12
MBA student	P13, P16
Engineering student	P15, P24
Ass’t project coordinator	P17
Corporate strategist	P19, P22, P23
Marketing, mgmt student	P20
Senior executive	P21

Figure 3.1. Participant Job/Role

3.1.2 Study Phases

This study consisted of four phases: 1) A diary task where readers logged all of their active reading activities, 2) a semi-structured interview about readers’ active reading behaviors and problems, 3) exposure to the then-current LiquidText prototype which I discuss further in the next chapter, and 4) a series of participatory design workshops where

Phase	Consisted of:	Served purpose:
Diary Study	• Users recorded instances of active reading. • Ran for 1 week.	• Gathered general knowledge about active reading. • Raised participants awareness of their active reading processes (for design workshop).
Interviews	• Enquired about current reading practice. • Sought intuitive hand gestures. • Evaluated prototype.	• Expand on diary study data. • Directly inform redesign of prototype. • Expanded participants' design and interaction vocabulary (for design workshop).
Design Workshops	• Two workshops, with 8 and 6 participants respectively. • Collaboratively reflected on current state of active reading. • Brainstormed new designs and features.	• Informed redesign of prototype to improve later evaluation. • Gathered additional reflections on problems in active reading.

Figure 3.2. Sequence of the three major phases of the formative study.

users collaboratively designed future active reading environments. In this section, I explain the methodology and reasons for each of these phases. The three data gathering phases of the study are summarized in Figure 3.2.

Diary Task

To provide a picture of real-world active reading tasks and behavior, I asked participants to record a diary entry each time they completed an active reading task over the course of generally one week (but up to three weeks). Each entry included answers to ten short questions addressing how, why, and where the task was performed, as well as the most difficult part of the task and how it could have been improved. I gave participants physical journals (see Figure 3.3) to act as tangible reminders to record entries, and we sent them daily reminder emails as well. Eighteen of the twenty participants turned in their diaries.

The diary task itself served two purposes. The first objective of the diary task was to understand the active reading process as it actually occurs, from readers in or near the situations where it occurs. This diary approach is consistent with earlier active reading studies [Adler, Gujar et al. 1998], and was intended to reduce our dependence on

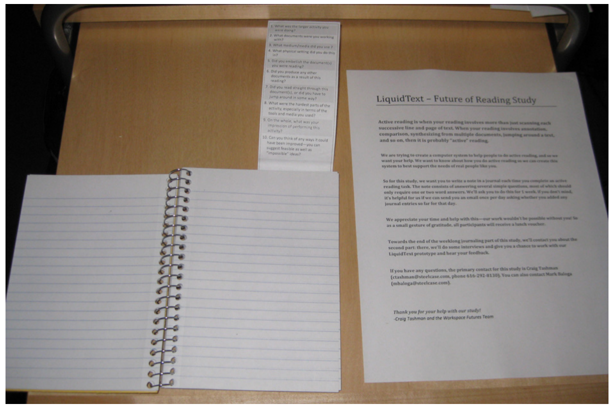


Figure 3.3. Materials given to participants for the diary task.

retrospective reconstructions of active reading events. The second objective was to make readers aware of their own active reading behaviors, difficulties, etc. This was to keep ideas and feedback in the subsequent phases of the study as closely grounded in the reader's actual reading situation as possible.

Generally, the diary study served two purposes: 1) obtain participant reflections on their active reading experiences while they were fresh in mind, and 2) make participants more aware of their active reading habits and preferences. The reason for #1 follows immediately from my first two research questions, above. The reason for #2 is because the last phase of the study—a participatory design workshop—requires participants to be experts in the domain of the thing being designed [Read, Gregory et al. 2002]; in this case, that would be active reading. But although people may do active reading regularly, if they have not recently reflected on the process, they may not have a good understanding of the actual tasks they perform or where the process breaks down.

Interview

After each participant completed the roughly weeklong diary task, I conducted one-on-one interviews with them. To the extent possible, I tried to schedule these interviews close to the end of the diary task so their memories of their active reading reflections would be fresh in their minds. The interview consisted of several parts:

This approximately 10 minute, semi-structured interview expanded on the themes in the diary study. I first asked participants about their immediate reflections on the journaling task itself, whether the study period was representative, etc. I proceeded to ask participants about their backgrounds, including demographic questions, their history with touch-screen devices, the kinds of reading they do, etc.

Principally, I focused on questions about active reading behavior. I asked about participants' current practices and the problems they experience in annotation, content extraction, etc. I also enquired if they had existing ideas about how active reading could be made easier.

Note that this discussion took place before participants[3] saw LiquidText. So, except for a small blurb about the system in the recruiting email, participants' views and ideas at this stage were not substantially influenced by my prior research.

LiquidText Evaluation

After completing the general active reading interview, each participant interacted with an early prototype of LiquidText. This component of the study, which took place during the interview session, included learning to use LiquidText, performing a representative task with the system, and an interview about their reactions. This phase of the study served two purposes. The first purpose was to elicit direct feedback on the experience of using LiquidText, which I discuss in more detail in the following chapter in the context of LiquidText's iterative design. The second purpose is rooted in my prior experience with participatory design, where I have generally observed participants creating incremental ideas, largely drawing upon the kinds of interactions they had already experienced. But an overarching purpose of this formative study, and particularly of the design workshops, was to learn how to improve the LiquidText prototype. Thus the second

[3] This is true for all but three participants, who were exposed to the system as part of their work.

purpose of the LiquidText Evaluation was to ensure the participants would be able draw upon the interaction approach it embodies if they thought doing so would be advantageous. But although this study is centered on LiquidText, I nonetheless wanted to identify participants' "uninfluenced" ideas. So to mitigate the risk of unduly influencing them, I waited until after both the diary task and the interview before exposing participants to the prototype (with the exception of P1, P7, and P22, who briefly saw the prototype prior to the study).

Design Workshop

The last phase of the formative study was the participatory design workshops. Here, participants collaboratively brainstormed about the state and problems of present-day active reading, before brainstorming about how a semi-idealized future would look instead. Each workshop lasted about 3 hours; and the day/evening before each workshop began, I asked each participant to re-read their active reading diary to refresh their memory of their active reading reflections.

After introductions, I began each workshop by asking participants to think about the current state of active reading. This began with 15 minutes of brainstorming the scenarios where people do active reading, and then proceeded with identifying the tasks within different active reading scenarios (such as highlighting, underlining, etc.) for 15 more minutes. Next, we spent 15 minutes considering what the biggest problems are in this active reading process. Although I explored the scenarios, tasks, and problems of active reading with each participant during the interviews, I raised these issues again during the design workshops so that participants could 1) better refresh their memories of these problems to be more prepared for the later parts of the workshop, and 2) build on the observations of other participants to provide possibly neglected details from their own experience. In case participants had difficulty starting the discussions, I brought lists of active reading scenarios, tasks, and problems taken from participants' reading diaries.

Also, all of the ideas participants thought of were written on post-it notes and stuck to a whiteboard.

Afterward, we began the redesign phases by spending 15 minutes brainstorming general, desirable characteristics for future active reading environments. Participants were instructed to think very freely, and be as abstract or concrete as they liked. They could think of addressing the problems they listed in the previous phase, or could think of unrelated, further out ideas.

Having brainstormed about general characteristics for future active reading environments, participants were then broken into groups of two to independently think of concrete design ideas for changing LiquidText. They were given up to about 10 minutes to do this, after which each group shared their ideas with everyone else.

At this point, the participants had created a list of less and more concrete design ideas, so I then asked each group to select several (about 3) ideas to expand upon. I asked them to make mock-ups and envision in-depth how one would use their designs. I gave each group a variety of raw materials to enable them to build mock-ups to reflect how one would interact with their designs. In particular, I gave them styrofoam and foam core boards of various sizes, paper, markers, transparent plastic to make overlays, tape, etc. To record their ideas, participants had laptops with webcams, which were configured to make flip-book animations that could be used to convey complex interaction sequences. Participants had about 45 minutes to design and construct these mockups, followed by about fifteen minutes of presentation. I closed each workshop with a brief, 10 minute discussion about which ideas participants thought were most important to implement in future versions of LiquidText.

Data Gathered

In total, I received 106 diary entries, with an average of 5.9 per participant (s.d.=3.4). As noted above, I took a liberal view of what constitutes AR, and accepted diary entries involving non-sequential navigation, annotation, the production of output documents or notes, the use of multiple documents at once, or content searching within or among documents. With these criteria, I excluded nine of the above entries. I also obtained 200 minutes of audio recordings from the interviews, and 6 hours of video form the workshops.

Analysis Methodology

Drawing from the diaries, interviews, and workshops I coded the data from each phase of the study independently, in order to allow me to distinguish in-situ data (the diaries) from retrospective data (the interviews), and both of those from collaboratively generated data (the workshops). I also coded the two workshops separately to allow me to identify differences between their results.

Within each separately coded part of the study, I principally relied on open coding as it readily identifies larger categories and themes in the data. In cases where many different dimensions of categorization were possible, I drew from prior AR literature (e.g., [O'Hara 1996]) in establishing first-level categorizations.

3.2 Findings

The results of this study include four parts. The first is a concise discussion of my analysis of the phenomena that occur within active reading, and the second is a broad discussion of how active reading is performed today; both of these are targeted mainly at RQ1. The third part details the problems that active readers face, and the fourth part describes readers' ideas for the future of active reading; these two parts principally address RQ2.

3.2.1 The Phenomena of Active Reading

In exploring current active reading behavior, I performed an open coding analysis of my participant interviews to draw out hierarchical categories and properties of the processes involved in reading. This would both support prior reading taxonomies (e.g., [O'Hara 1996]) as well as provide a standard against which to compare LiquidText's breadth and depth of features.

The analysis of this data resulted in a detailed picture of the numerous phenomena within the active reading process. At the highest level, it yielded several overarching categories: *annotation, manipulation, navigation, comparison, visualization/layout, creating separate content,* and *abstract challenges*. Within each of these are sometimes several levels of sub-categories, describing things such as the purposes for which people perform a type of action, or the specific forms that action can take. Many of the (sub)categories also have dimensions and properties along which their constituent phenomena vary.

A striking result of this analytical approach, brought out by the categories and dimensions, is the level of variety spanned by different tasks within active reading. For example, annotation that is intended to support re-finding ranges from providing a local cue (like highlighting) to a global cue (like bookmarks); from supporting retrospective memory to prospective memory; from supporting retrieval of the annotations themselves to retrieving content in the original text, and more. Similarly, navigation-based search may vary in terms of the specificity as well as quantity of the results being sought. Also, the area to which content is copied (or in which it is created) during active reading may vary from being a local to a global repository; a temporary to a long-term repository; or formal to informal/ad-hoc in structure. There are a variety of other examples as well, all of which help to inform the design of systems like LiquidText by providing a detailed picture of the breadth of requirements for the different aspects of active reading.

Cumulatively, this description of the processes within active reading served as a framework for understanding active reading throughout this research. It provided several of the insights into how active reading is performed today, as described below, and generally provided a benchmark for assessing the feature-set of LiquidText.

3.2.2 The State of Active Reading

Drawing heavily on the above categorization, as well as other material provided through participant's diaries and the design workshops, I present here a broad picture of the way active reading is performed today. While several active reading challenges are found in this section, that topic is principally addressed in the section that follows, with this section focusing on the behaviors of reading.

Reading goals: As a step toward ultimately understanding the problems faced in active reading, I first sought to understand readers' motivations for engaging in these types of tasks. My findings here came principally from the diaries and were in mostly line with existing literature (e.g., learning, condensing, etc. [O'Hara 1996; Murray 2003]). However, I did identify two additional goals. The first, *organizing content*, was observed 11 times (out of 63 unambiguous entries), such as when P2 was clustering her interview transcripts by topic, or P3 was clustering car repairs by priority. The second, *comparison*, was observed twice, such as when P15 was evaluating different configurations while shopping for a personal computer.

Personal reading: This shopping example is also significant as an example of personal active reading. Although active reading is generally taken to be a work-related phenomenon (e.g., [Sellen and Harper 1997; Adler, Gujar et al. 1998]), participants revealed four examples to the contrary, including P17's highlighting key quotes from a novel, or P1's claim to annotate virtually everything he reads except fiction. And expectedly, personal active reading tended to be performed in personal, informal spaces, including lawn chairs and couches.

Active reading locations: While the importance of mobile active reading is known [Adler, Gujar et al. 1998], I was surprised by the breadth of mobile settings used. For example, I found participants doing work-related active reading in bed, at picnic tables, at the kitchen table, and in a car. Generally, of the 73 diary entries with unambiguous locations, 19 (or 26%) involved active reading in mobile locations. So, in line with earlier work [Adler, Gujar et al. 1998], this suggests readers broadly require a relatively portable means to support active reading.

Computers and paper: Complicating the need for portability was readers' continued reliance on multiple media for performing active reading tasks. Similarly to earlier work [Sellen and Harper 1997], this study found that approximately 25% of unambiguous diary entries described using paper-plus-computers for a given task. But while [Sellen and Harper 1997] found that a majority of the time, document tasks were paper-only, I found now 63% of active reading diary entries were exclusively performed on the computer—suggesting that computers are indeed overtaking paper as the active reading medium of necessity. This is not to say that it is the medium of choice—participants generally continued to express a preference for paper in active reading tasks. Even the extremely paper-like tablet-PC was described by P21 as requiring too much time to access the inking functionality.

One reason people mixed computers and paper was corporate policy—as in filling out an online expense form. But in other cases, computers were simply better at handling some parts of a task. For example, P23 described printing documents to read them, before going back to the computer to highlight them. Generally, when comparing media, tasks demanding more "pure" reading were seen as better suited to paper. Part of the reason for this was that paper could show more text at once than a computer.

Significantly, this shortcoming of computers was noted in 1997 [O'Hara and Sellen 1997], but was suspected to have been resolved by 2007 [Morris, Brush et al. 2007]. And while the benefits of larger displays have been seen for specific active reading tasks under laboratory settings [Costabile, Paternò et al. 2005], we find that it is still a problem in actual active reading as well. And the most obvious solution—increased screen space—conflicts with the importance of portability.

Besides display space, others described a preference for the tangibility and navigability that paper afforded; one further noted how that physicality can lead to feelings of guilt over printing to paper.

But paper had drawbacks as well. P12 and P6 both described frustration with physical books because of the difficulty reading near the crease of the page, and the weight of many hard-covers. In line with prior lab studies [Pearson, Buchanan et al. 2009], P1 and P16 discussed paper's lack of margin space, while others felt entering/editing information is harder using paper than computers.

In a similar manner to [Sellen and Harper 1997], participants in this study saw computers as perhaps better suited to linear, structured AR tasks. P2 described them positively as, "forcing me to be neater." But paper appeared superior for tasks with many parallel components, without clear structure, and where added flexibility was needed. P2 continued,

> *"When I work in digital form, I want to be more codified in how I approach things...with pencil and paper, I like the ability to feel loose. It doesn't feel like as much of a process as it does when it's ...on the computer."*

But as computers and paper each have domains where they are superior, some research has found people comfortably using them together (e.g., [Marshall, Price et al. 2001]). But by contrast, my participants revealed considerable frustration using paper and computers together. Part of the problem was that readers felt frustrated that paper does not support computer-like functions like drag-and-drop. The difference in spatial orientation caused regular frustration as well when participants had to switch from looking up at their monitors to down at their paper. Part of this disparity may be resolved by putting computer monitors and paper at a common viewing angle, but the differences in contrast and light reflection properties noted by participants may make a solution less forthcoming.

Output materials: Active reading is known to involve collaboration in certain domains [Sellen and Harper 1997], and we found that over 73% of the 57 diary entries with clear outputs (i.e., materials resulting from the active reading task) were collaborative. And while the outputs varied between broadcast (e.g., blogs) or narrowcast (e.g., presentations), and in creating content new or modifying existing material, the single most common type of output was email, which was found as an input and an output in about 20% of entries. By contrast, earlier studies pointed to email as a way to send documents used in active reading [Sellen and Harper 1997], but did not suggest emails themselves as the focus of active reading tasks. Besides email, we found a variety of other collaborative materials were produced, including summaries, analysis reports, PowerPoint slides, etc. Generally, we found that active reading is not confined to documents designed for careful, deep reading. P21, for example, describes trying to make sense of all the different authors involved in several long chains of email.

By contrast, 27% of outputs were not intended for sharing, but these were usually intermediate results, such as lists of notes, outlines, or sketches.

3.2.3 Active Reading Problems

Readers listed numerous active reading breakdowns related to many of the process's major phenomena, such as navigation, annotation, search, content awareness, etc. Most of these problems came from the diaries, where I asked readers about the hardest parts of their reading experience. Analyzing the results yielded 92 clear issues, which I organized into 21 categories. I also noted 30 issues identified in the design workshops, and several more in the interviews.

Some of the most prominent categories of challenges are partially discussed in the above section. Visibility, for example, comprised 11% of issues identified in the diaries, and included seeing many lines of text, documents or notes at once. Media switching, between paper and computers, comprised about 8%. In this section I discuss several additional categories of issues.

Navigation: Inter-document navigation issues constituted over 16% of those identified in the diaries. One well-known example of this is switching among many application windows (and, as I found, spreadsheet tabs) [O'Hara and Sellen 1997] involved in an active reading task. This was partly because readers forgot their places in their documents, as P19 described this in his diary,

> *"[It is hard]...jumping back and forth between documents and trying to remember where you were. I would do this much differently physically. I'd have my five documents spread on the table and make a mark where I was at."*

Likewise, readers struggled with just keeping track of all the different materials and their significance. P13 explained this, "[It's hard] managing opening multiple documents, understanding the context, relevance, and usage of the documents." This difficulty, as well as the regularity, of multi-document scenarios raises particular challenges for designers. It puts even greater demands on already limited visual real-estate, perhaps demanding more creative visualizations or hardware form factors as solutions.

Besides inter-document navigation, about 5% of issues identified related to *intra*-document navigation. For example, P16 notes, "…especially when the print-outs are not on the same page; printed on different pages and you've got to flip back and forth which drives me nuts."

While prior research described how readers freely arranged pages of paper documents to view disparate document areas in parallel [O'Hara and Sellen 1997], this clearly breaks down for bound documents, as the quote illustrates. In some ways then, intra-document paper navigation is similar to inter-document navigation on computers: one can often see only one page at a time and so must rely more on working memory than perception to integrate disparate content. That is, paper's spatial flexibility is important to readers, allowing them to arrange things freely, but this seems to be where that flexibility reaches its limit.

Generally, one implication of readers' navigational difficulties is to emphasize the point that AR navigation is non-linear. Readers do not consume a document or a page and move on to the next one. Rather, they cycle between documents and portions of documents often in complex integration tasks.

Annotation: In the diaries, annotation issues were less common, comprising about 5%, and largely relating to software limitations, such as applications which did not allow comments or highlights, forcing the reader to put thoughts in separate documents. The workshops and interviews though, revealed several additional issues. In the interviews 5 readers cited difficulties making margin notes on paper, especially due to the lack of physical space for doing so. Other comments concerned managing annotations—such as re-finding one's annotations after the fact, a desire for easier editing of sticky notes, and a desire for annotations to intelligently make themselves known at the times and places where they are needed.

In an interesting counterpoint to readers who found themselves unable to annotate at all, one of the workshop discussions considered the opposite—the temptation to annotate too much. This effectively makes things like highlights meaningless, leaving the reader uncertain of why she annotated that material. This may partly be due to readers' use of media that cannot support the full breadth of different annotations (with concomitant semantics) that readers produce.

While annotations have been studied before [Renear, DeRose et al. 1999], my

	Simple semiotics	Generative semiotics
Inter-page	*Dog-ear*	*[unknown]*
Intra-page	*Underline*	*Sketching*

Figure 3.4: Two dimensions of non-text annotation, with examples for each quadrant.

analysis suggests that non-text annotations can be categorized along two dimensions (Figure 3.4): *inter-page* (i.e., marks visible across pages, such as a dog-ear or bookmark) versus *intra-page* (like an underline), and *generative semiotics* (like sketches, where the reader can put marks together to construct more complex meanings) versus *simple semiotics* (again, like underlining). Different media then, support different points along these two dimensions. For example, a traditional word processor supports simple, intra-page semiotics like highlights and underlines, but makes it awkward to create marks complex enough to be generative. Likewise, a book supports intra-page as well as inter-page, simple semiotics, such as dog-earing and sticky notes. And notably, while books also support generative semiotics within the page, no commonly used medium seems to offer support for semantically rich markings that can cross page boundaries (inter-page, generative semiotics). So the use of media that do not adequately support this space may be a way to look at why readers could not express their annotations in a way that would be richer and thus more memorable.

Awareness: One challenging aspect of active reading was maintaining an awareness of all of the background information relevant to what one was reading at the moment, which comprised about 8% of issues in the diaries. This information included context established elsewhere in the documents at hand (such as who said what in an email), as well as one's earlier thoughts and reflections, and even location within the material. For example, P21 described doing frequent flipping through a document because of the difficulty of trying to keep so much data in one's head at once. By contrast, P24 described taking notes from multiple sources but losing track of where each note came from. This latter issue is interesting as it highlights a challenge of using paper in multi-document work, as it appears to offer little support for capturing or viewing cross-document relationships.

Note taking: While the diaries did not include challenges relating to note taking, several were mentioned in the interviews. Many of the issues cited came from creating new notes, ranging from the mechanics of the copying text (P23), building and managing outlines (P11), and anticipating what will be important in the context of a text in order to decide what to copy (P2).

Participants also described issues such as determining how best to organize their notes, or how to integrate together notes from multiple sources. And, because notes are separate from the source text, a document switching task was often involved as well. This was explained by P2 during the design workshops:

"[Annotation] gets noisy because I'm highlighting or tagging different things in the same document; and [with notes] I'm creating new documents but then I've got a slew of new documents and I'm constantly going back between the originals and the new documents…"

Notes then, in contrast to annotations, appear to be at a greater risk of reducing the efficiency of the reading process. Nonetheless, their function as a means of centralization and aggregation appears to be essential.

Retrieval: Even after their creation, readers continued to struggle with notes—and annotations—in the retrieval process. Like annotations, retrieval of notes was hard in part because readers needed a thought or excerpt about a document at a particular moment in life when it was relevant, and it was hard to make these connections naturally. As I discuss below, readers mentioned some ideas to address this, including integrating an active reading system with a to-do list.

The design workshops also included considerable discussion on the problem of recalling needed context when returning to an entire AR task. This observation led to several participant ideas discussed in detail in the next section.

In summary, in addressing RQ2, participants in this formative study pointed out a variety of ways that active reading can be challenging. First, they showed clearly that paper and computers each have shortcomings—paper, for example, can be flexible but often laborious, while computers can be efficient but in some cases overly structured. Among the major problem areas we considered (navigation, annotation, awareness, note taking and retrieval) one frequent theme was re-acquiring needed mental state—as in recalling the meanings of annotations or notes, or even of entire documents. This difficulty has led multi-document AR tasks to be especially challenging because of the increased memory burden, but even single-document tasks can have this problem when the text is a large document. Mechanical issues are less common, but still significant when creating notes. In the next section, I discuss our participants' ideas and solutions for addressing these and other AR problems.

3.2.4 Participant Ideas

In this section, I present the ideas that participants suggested for the future of active reading. Throughout the study, participants generated over 170 "ideas," ranging from vague suggestions and incremental changes to radical, novel ideas. These ideas also ranged from addressing active reading itself (such as ways to annotate) to tangential tasks (such as to-do list functionality). And consequently, while many of these ideas could be relevant to LiquidText, many others are not. I discuss all of these however, because they offer insight into the types of tools that people want integrated into their active reading environments, and thus have bearing on the design of future active reading technology beyond LiquidText.

Sequentially, I begin by presenting ideas generated in the diary phase, before participants were exposed to the LiquidText prototype, followed by those coming afterward, during the design workshops.

Diary Ideas

The most situated of participants' ideas came from the diaries—where participants focused on improving their immediate active reading experience, as opposed to supporting entirely new experiences. I classified the 65 relevant diary ideas into 14 themes; the most prominent of which are presented here.

Visual layout of content: Likely in response to the significant problems in navigation and visualization, this category focused often on getting disparate content into visual proximity. Suggestions sometimes took the form of changes to existing applications, such as P19, who wanted Excel to freeze more than one column at a time and PowerPoint to show more than one slide at a time. Other suggestions included ways to view all the highlighted portions of a document at once, or various visualization ideas such as P13's using visual links to show document relationships, or P5's suggestion to use focus-plus-context views.

Increased directness: Several participants described using more direct ways to interact with, and especially copy, text. While one offered a pen-based approach, most were interested in touch-based methods to drag content, as with P21,

> *"Would be great for touch functionality…just touch the screen to highlight something, flip between pages, etc.; That'd be much easier than dealing with the tablet pen."*

It is important to note that our recruitment instruments mentioned a touch-based prototype reading system, so even before seeing the prototype itself, our sample might have been especially inclined to the idea. Still, their ideas are worth noting as ways touch could be used in support of active reading.

Linking: Also likely as a response to navigation problems, several participants described improved ways of linking pieces of content together. P3 described "threads" that could connect paper documents together, and then extend to link to digital material as well. Others described linking audio notes to a location in a text, or a centralized log book linking notes to many different documents.

Annotation expressiveness: Potentially as a response to the above problems of inadequate annotation semantics, several readers also described richer ways to embellish content, such as more colors of markers or color-changing paper. Likewise, P9 described dimming irrelevant parts of a document as a way to reduce visual saliency—effectively the opposite of a highlighter, which increases saliency. Participants also suggested multimodal annotation, in which voice could be used to annotate content.

Collaboration: In a reflection of the often collaborative nature of the inputs and outputs of the AR process, three participants described synchronous and asynchronous collaboration ideas. For example, P21 wanted long emails to be color-coded to show what text was written by whom.

Workshop Discussions

Unlike the diaries, the workshop discussions afforded participants an opportunity to build on one another's ideas, allowing them to focus on developing the thoughts that were more interesting to the group. While many individual ideas were raised, here I discuss several of the themes that come out of this group discussion.

Collaboration: Participants in both workshops independently showed a strong interest in collaborative active reading, both synchronous and asynchronous. The scale of sharing, though, varied between quick excerpts and one's entire active reading workspace—suggesting such variability would be important to capture in an actual collaborative active reading system.

State capture: Both workshops also discussed a desire to capture the state of AR environment. Workshop 1 (WS1) was mainly interested in recording one's AR history in order to build a rich corpus of cues to help readers regain their train of thought when returning to an AR task after an interruption. The participants suggested several sources for these cues, including the paths one had taken through their documents, the meaning of the documents so far, and events going on in the larger context/world of the reader.

WS2, by contrast, explored state capture as a way to switch between and transport "workspaces" of content. In particular, participants imagined a large display surface with many documents open, which could be "minimized." Later, the user could return to it or to another workspace. Participants went on to envision a special "magic pen" that, besides being used for annotation, could act as a proxy that could store all or part of a workspace. The user could then bring this stored space to a colleague to support collaboration.

Hardware platform: Developing a large-screen workspace, participants in WS2 considered several details about their hardware platform. First, participants expected a multitouch-plus-stylus work surface. Touch was seen as better for manipulation, like scaling and moving materials, whereas the stylus was seen as better for selection and highlighting. The size of the system was also important; participants wanted to see everything at once—likely a response to the visibility problems discussed above. Participants also wanted this large display to switch between upright (and curved) and flat, as there was interest in both positions.

Addressing the cross-media problems described above, participants considered embedding scanners in their workspace to easily digitize papers (as in [Sun Microsystems 1992]), which they refined into a paper-sized tablet computer sitting flush in a large table-top surface. The tablet could then be removed to focus on one document in a different setting (e.g., at a café), and returned to the table-top when the user needed more space. We found this an interesting resolution to the paper-computer dichotomy: replacing both media with a unified hybrid, especially since modern tablets do appear comparable to paper [Morris, Brush et al. 2007].

WS1 focused less on hardware platforms than WS2, but did explore input/output options, like haptic feedback, lights, and sounds as cues to events in an active reading system. In an effort to let the user absorb more material, they suggested parallelism: having the system read annotations aloud, while the user read the document text.

For input, WS1 participants were interested in again exploiting parallelism. They envisioned multiple input modalities, including the use of gaze and chorded keyboards (which they borrowed from pianos) for tasks like highlighting, and a stylus for creating and manipulating content. As with output, participants seemed interested in maximizing efficiency, but also in having what they called “natural interactions” for intuitively controlling their active reading environment

Sense-making: Of particular interest to WS1 were ideas to help readers better understand a text. Visualization played an important role in this, being used to show relationships among documents and especially to give readers a *sense* of a document without requiring them to read it. Ideas included semantic zooming, to show in-place keywords for a given region of text, and “tag clouds” for terms in the document. Recognizing that many of their documents may have been read and reviewed by others, participants suggested automatically searching the web for summaries and presenting tag clouds of their most frequent terms. To help readers identify important areas of a text, WS1 also suggested readers

But once the reader decides to look at a text in more depth, participants also wanted to help her identify the most relevant areas. One approach was to do this entirely automatically; another was for the reader to tag some portion of the text as important, then the system would identify other text that is similar—either at the word or conceptual level.

Annotation: In line with the interest in more flexible annotations noted in the diaries, WS1 considered ways to annotate annotations. They discussed organizing annotations into layers, and assigning annotations levels of importance. They also wanted to use these rich annotations as a search mechanism, allowing the user to find all parts of a text that have been tagged a given way.

Rather than add more structure, WS2 was interested in freeform annotation. They envisioned, on their large, pen-touch workspace, the reader could freely write not only on documents, but across them, depicting more complex inter-document relationships. This fits well with the importance of inter-document scenarios.

Organization: Participants differed in the level of organization they desired, as evidenced by an outright disagreement in one workshop about whether excerpts of content should be organized in well-structured folders or in freeform piles. The result is instructive—though many readers may need organizational affordances, the amount of structure these impose needs to vary depending on the person and the situation.

Workshop Mockups

Having brainstormed an array of ideas, the prototype mockup activity gave participants a chance to integrate and refine a selection of those ideas into simple system designs. Here, I review the general direction taken by several groups, as well as the ways their designs built on their earlier ideas.

WS1, team 1: Participants here focused on recording and presenting a history of one's active reading process. They envisioned a timeline, which one could pan/zoom using multitouch gestures, shown in parallel with the notes one was taking at a selected time. To provide context, the timeline would be populated by one's Outlook calendar, and would contain pointers for when they began reading different texts.

WS1, team 2: Team 2 stressed tangibility and portability, designing a tabletop plus e-paper hybrid environment to allow the user a seamless transition between the two media. These two systems would be synchronized, ensuring that notes taken on one would be reflected on the other. Taking tangibility even further, they included e-paper post-it notes which could be attached to the e-paper document, and whose positions/content would also be reflected in the tabletop system. They could then can be linked and labeled.

WS1, team 3: Like team 2, team 3 built on a tabletop plus e-paper hybrid, principally using the e-paper as a more comfortable way to input text. They expanded upon it by letting the user write notes and sketches, which would be associated with whatever else the user was doing at the time (e.g., meetings, etc.), as well as by exploring visualizations for showing relationships that cross documents.

WS2, team 1: Akin to what WS1 teams devised, these participants built on a tabletop-plus-tablet hybrid. In addition to their earlier ideas of freeform annotation across multiple documents, they proposed a means of aggregating notes, as well as accessing common tools like calculators, one's calendar, and to-do list. The latter of these also functioned as a workspace manager, building on the state management ideas discussed in the brainstorming. This allows one to click a to-do item to move to the associated workspace.

WS2, team 2. The second team envisioned a rich system to support collaborative design activities—with active reading as just one piece of the process. It included an upright monitor used for face-to-face video conferencing, as well as shared and private workspace areas for content creation, reading, sketching, etc. Integrating active reading into their larger workflow, it allows complex diagrams to be linked to different points in documents. And as with other teams' ideas, this too included full history recording and an integrated to-do list, where content could be moved for later reference.

WS2, team 3: This team also looked at supporting individual or collaborative active reading, but focused more on collocated scenarios. Like other teams, it also included upright and flat display surfaces, allowing the user to drag documents between them. It supported document linking to create structure, but allowed the user to push all documents to the side to make a visually clean workspace as well.

3.3 Answering the Research Questions

One of the core goals of this formative study was to address my first two research questions, which again concerned 1) the circumstances and tasks of active reading, and 2) the challenges readers face, respectively. In this section, I begin by summarizing the above findings and highlighting how they address these research questions. I conclude by explaining the implications for the design of LiquidText—which was also a goal of this study.

3.3.1 RQ1: Active Reading Practice

As introduced in Chapter 1, the first research question concerned the active reading process: **What are the circumstances surrounding active reading as it stands today, and what lower-level tasks do people perform as part of the active reading process?** Throughout this chapter, I have sought to address this through a categorical analysis of the phenomena that occur within active reading, as well as by generally elucidating and updating our understanding of the manner and circumstances associated with the reading process.

In seeking to identify the phenomena that occur within active reading, as well as the overall goals of the readers, my results were largely in keeping with the earlier understandings that I detailed in Chapter 2 (e.g., [O'Hara 1996]). I did however identify two overarching goals not regularly described in the prior literature, including *organizing content* and *comparison*. While both of these are known to be steps within other reading tasks, they are not typically discussed as goals in themselves. Identifying them as such raises the importance of explicitly and comprehensively supporting them within active reading.

In the broader context, the study highlighted the relatively recent prominence of computer-mediated active reading, which constituted 63% of participants' diary entries—plus an additional 25% for computers used in conjunction with paper. Accompanying this rise in computer mediation is the variety of computing contexts involved. Participants described active reading of emails, as well as spreadsheets and web forums. And the outputs resulting from their reading were likewise often computer mediated, including PowerPoint slides and especially email. In seeking to better support active reading technologies, this is significant for several reasons. First, is shows that improving the computer-mediated active reading experience is relevant, as it has become vastly more prevalent than in earlier studies [Sellen and Harper 1997]. And second, it highlights the point, alluded to by the ideas developed in the design workshops, that active reading is not really a separate task that one performs independently of other work. Thus, while it is outside the scope of LiquidText, a future research direction could include exploring how to design computing environments to support active reading as a whole, rather than just in select, specialized applications.

The circumstances of active reading were also surprising for their breadth. Participants, for example, routinely described performing active reading tasks in highly transient settings, such as kitchen tables, cafés, or beds. This is consistent with earlier work that found active reading taking place away from one's office [Adler, Gujar et al. 1998], but goes further, showing clearly the widely varying levels of control and transience with which participants must cope. For example, one's kitchen table and a café are both mobile settings, but come with very different expectations of how long one may remain there, and the extent to which one may modify the setting, say, by moving the furniture. For designers of active reading technology, this helps draw attention to the breadth of settings a user may occupy—in terms of ergonomics, lighting, social expectations, persistence, and so forth.

But while active reading is in some ways broad and adaptive, it can also be rather conservative. Of all the diary entries, only one included an example of new, social media. This was P11, who was creating a blog entry at the time which involved active reading. Similarly, only P1 performed active reading using a small mobile device (a Blackberry). Generally, it seems active reading tasks and habits may evolve somewhat more conservatively than computing as a whole—perhaps because the software, as well as the users, need time to adapt to handling the rigors of active reading on such platforms. Notably though, I conducted this study shortly before the iPad became available, so it is possible that people are attempting to perform active reading on that platform more frequently.

3.3.2 RQ2: Active Reading Problems

In contrast with the descriptive nature of my first research question, RQ2 was more prescriptive, asking how reading should change, specifically: **What difficulties do people tend to experience during active reading? Particularly, which aspects of active reading are most challenging?** Some of my findings here echoed those of earlier research (e.g., [O'Hara and Sellen 1997]), particularly the many difficulties of reading with computers. Screen size, for example, was sometimes still too small to view an adequate amount of text. Going beyond prior literature, my findings also raised other problems, especially with paper. Participants discussed the lack of margin space it affords, as well as the difficulty reading near the crease of a bound book. And while physicality was one of the benefits of paper, it also led to objectionable size and weight in some cases. Given the shortcomings respective to computers and paper, participants often used both media in tandem for a given task, which led to considerable frustration. Together, these problems highlight that there is still no “silver bullet” for active reading. It is thus counterproductive to seek to simply mirror the experience of one medium in the other (as with computers that emulate paper).

And beyond the scope of specific media, we see that a variety of areas of active reading were challenging for participants on both paper and PCs. As noted above, participants struggled often with visualization, such as in being able to see sufficient lines of text, documents, or notes in parallel. Similar to visualization was awareness, where participants struggled in maintaining a sense of all of the background material relevant to their task. This included other related texts, as well as one’s own earlier thoughts and reflections. I argue that these difficulties are suggestive of underlying problems in the way that content is represented to the user—the user cannot easily view the quantity or type of content required by the task. As a result, and as discussed in Chapter 1, LiquidText focuses strongly on providing a more flexible, controllable visual representation of text.

Beyond just the representation of text, participants' difficulties show opportunities for improvement in interaction as well. Navigation was a prime example of this, especially when working with multiple documents. The mechanics of changing from one document to another, using documents in parallel, and simply keeping track of where one was in a document all wanted for improvement. Navigation within a single document incurred problems as well, as when repeatedly flipping between pages of a bound text. Likewise with annotation, participants described challenges in re-finding highlights or notes, as well as simply organizing them. There were also cases where annotations were not rich or varied enough for the task, and eventually become meaningless through overuse. Together, these problems suggest a lack of control as well, not over the static representation of content itself, but in the dynamics for changing and building on the representation. Thus, LiquidText seeks to offer more control and flexibility through richer visualizations as well as tools to manipulate those visualizations.

3.4 Application to LiquidText

Cumulatively, the results of this formative study could give rise to numerous reading technologies, ranging from new hardware platforms to collaboration services. Ultimately though, only some of these were applied to LiquidText for two reasons. First any added functionality should contribute to LiquidText's embodiment of the notion of flexible, high degree-of-freedom representations I discuss above (or otherwise make LiquidText more usable for evaluation), as understanding such representations is the objective of this research. Second, any added features must be technically feasible in the hardware and software framework I have available. The specific changes derived from these inquiries into the status, problems, and ideas associated with active reading include:

- Annotation aggregation: one of the problems participants discussed with current active reading media is retrieving annotations; as such, I added functionality to LiquidText for supporting visual annotation aggregation.

- Arbitrary linking: participants' designs for future active reading technology often incorporated a notion of building links between pieces of text flexibly and easily. This fits well with the goals of LiquidText, and so a variant was included in the system.
- Annotation metadata: In the workshops, participants' were interested in being able to layer their annotations so they could better organize them, or select which ones to view. This fits with the goals of LiquidText as it gives the user more dimensions of control over how data is organized, and so a simple form of this functionality was added to LiquidText. But as described below, it was disabled in the final system.
- Freeform annotation: participants' comments and workshop mockups made clear that freeform, drawn annotation is an important part of active reading—and as it provides more flexible ways of imputing annotations, it was well in line with the goals of LiquidText. A version of this was thus incorporated into the prototype, but was ultimately disabled for the final system, for reasons described below.
- Overviews: Workshop participants described several visualizations for better gaining an overview of a text and its relationships. Thus, since it too would give the user more control of the visual arrangement of content, I added a form of focus-plus-context overview visualization to the prototype.

In addition to these, various other feature additions and changes were made based on the direct evaluation of LiquidText that was also a part of this formative study. This, however, is described in the following chapters, which describe the LiquidText system and explain its design process in detail.

CHAPTER IV
LIQUIDTEXT DESIGN PROCESS

In order to investigate the latter three research questions I discussed in Chapter 1, I developed a high fidelity prototype software application: LiquidText. The application aims to be a usable embodiment of the abstract notions of flexible, high degree-of-freedom representations discussed above. Likewise, it operates on a multitouch platform.

In this chapter, I provide an outline of the processes and methodologies through which I designed and developed LiquidText, beginning with a discussion of some of the basic design decisions that went into it—this will be a review of, and expansion upon, arguments made in Chapter 1. Afterward, I will discuss the general design process for the system. This chapter, though, will remain fairly abstract; in the following chapter I will explain the concrete functionality of the system, and how it evolved to its final form.

4.1 Basic Design Choices

Underlying the choices of specific features for LiquidText are two fundamental design decisions that support the rest of the system. The first is the choice to use a high degree-of-freedom, non-paper-like document representation. The second is the use of a multitouch interface. I provide the rationale for both of these decisions in the introductory chapter, but review them briefly here, and expand on my rationale for the use of multitouch in light of the discussions in Chapter 2.

4.1.1 Representational Approach

As I argued in Chapter 1, my use of a flexible, high degree-of-freedom approach to document representation stems from several shortcomings of paper and traditional paper-like computer interfaces. In chapter 1, I pointed out that several problems are rooted in paper's representational fixedness. In brief summary, these include:

- Difficulty comparing disparate parts of a document.

- Difficulty maintaining original context and order when doing comparisons.
- Poor support for document overviews.
- Little support for ad-hoc navigational structures.
- Difficulty making annotations that refer to disparate text areas.
- Bounded scale of annotations on the page.
- Difficulty retrieving annotations

The active reading design requirements I discussed in the chapter 2 suggest further points for this list. An important aspect of annotation, for example, is the ability to find and navigate among the annotations [Renear, DeRose et al. 1999]. However paper offers little support for working with annotations independently of the document itself, especially while keeping them visually in context. This is similar for extracted content, which likewise requires a close relationship to the context from which it was taken [O'Hara, Smith et al. 1998; Marshall, Price et al. 2001]. These theoretical complaints are supported by the empirical investigation of the previous chapter, where participants specifically noted the difficulty of re-finding and maintaining awareness of annotations, and of remembering the relationships among excerpted content.

Paper also provides a fairly weak sense of location within a text. Although one can readily judge the overall percentage of completion (by showing the breadth of pages that have been read), a more detailed picture of the reader’s position relative to the logical structure of the text is less forthcoming.

As I argue in chapter 1, many of these and other issues can be traced, at least partly, to paper’s representational inflexibility. For example, by merely relaxing the requirement that the spatial structure of a page be fixed, one can consider ways to bring disparate pieces of text together, alleviating difficulties comparing or annotating distant areas of a document. Similarly, one can distort a text so as pull comments and annotations away and together, viewing them in aggregate but without breaking their link to the original text. By admitting varying scaling factors, one can use semantic zooming to show annotations in context, while miniaturizing intervening text. And similar techniques could be used to provide needed overviews of a text, while still displaying detail where it is required. And though the current prototype does not now do all of these things, a general, flexible representation in this vein is what LiquidText seeks to provide. However, introducing this degree of flexibility, with fewer constraints on the visual/spatial properties of a document, can result in a representation that is difficult to control efficiently. This is part of the reason for the second key design choice in LiquidText: multitouch.

4.1.2 Input Approach

As I describe above, LiquidText is designed around the use of multitouch and gesture-based input. There are two principle reasons for this: 1) the specific affordances of LiquidText's high degree-of-freedom representation benefit from it, and 2) active reading generally, regardless of the medium, benefits from it.

The reason that active reading generally lends itself to multitouch input is because this reading process involves parallelism, particularly in terms of parallel visualization and parallel manipulation. Based on O'Hara's reading taxonomy and O'Hara and Sellen's media comparison, Figure 4.1 helps indicate which active reading goals involve these forms of parallelism [O'Hara 1996; O'Hara and Sellen 1997]. We can see that reading to learn, reading for research, and reading for critical review all involve parallel viewing of documents. And if multiple documents are visible, I believe it is reasonable for a reader to seek to manipulate these documents in parallel as well. Additionally, reading for research, reading for summarization, and reading to search are explicitly understood to involve some level of parallel document interaction. The requirement for parallel interaction is thus found in a variety of types of active reading. That, combined with other research emphasizing the importance of parallel, bimanual interaction with texts [Morris, Brush et al. 2007], strongly suggests the appropriateness of a multitouch input device, as it wound readily support parallel, bimanual interaction.

Reading Goal	***Tasks included in reading goal***
Reading to learn	Brief, skim-based reviews for rereading. Viewing documents, or parts of documents, in parallel. Smooth navigation between sections. Outlining/note taking. Underlining.
Reading to search or answer a question	Keyword scan, followed by reading in detail. Find and shift between bits of info, to make inferences. Zig-zag reading through text as search strategy. Reading alternate sentences as a search strategy. Hold finger at a page containing important content. Developing a sense of the organization of a document.
Reading for research	Working with multiple documents. Scanning/skimming to decide whether to proceed. Non-serial scan of document to get relevant info. Integration of material from multiple documents. Note taking/outlining.
Reading to summarize	Record main points of text, or a plan for the summary. Annotation or note taking. Idiosyncratic annotation for re-representing document structure. Moving and within between documents during task. Navigating multiple documents in parallel.
Proofreading	Relocating exit point when re-entering document.
Reading while writing from multiple sources	Formation of intermediate texts for planning, reviewing, reshaping source content and reader's thoughts. Continual movement between and within documents. Scanning texts to find related and relevant info. Developing a sense of text, required for efficient navigation. Develop sense of document space. Reading own compositions.
Reading for critical review	Moving between parts of text for comparison. Testing consistency of ideas across sections. Obtain implications of results (test consistency between sections, review methodology, assess data interpretation, review several types of data concurrently) Develop good sense of text. Moving between different sections. Laying out several sheets at once for quick scanning or simultaneous viewing. Annotation and note taking.
Reading to apply	Re-accessing of text at appropriate points in document.

Figure 4.1. Media-related tasks entailed by different active reading goals.

Of course, other active reading systems have included some level of parallel, bimanual interaction without using multitouch. However, these systems have generally used the non-dominant hand only to advance through the text [Schilit, Golovchinsky et al. 1998; Morris, Brush et al. 2007], and as such accept only very simple input from the non-dominant hand. In the vernacular introduced in chapter 2, they offer support principally for only one type of parallel-cooperative-*asymmetric* interaction between the hands; that is, one hand sets the reference frame, and the other does the work within the frame. But is this adequate? Parallel-cooperative-*symmetric* interaction, for example, could be important to support as well, such as when a reader wants to position two documents at once to view side by side. *Parallel-independent* interaction can likewise play a role, such as when a reader uses a thumb to save a place in a text while using the other hand to flip the pages. Effectively, paper offers more opportunities to take advantage of the efficiencies of bimanual input than most computer-based active reading media. Thus, by not providing support for the categories of bimanual interactions that are possible with paper, we are potentially putting our computer-based active reading media at a fundamental disadvantage. In order to avoid this, and to have the flexibility to design a wide range of bimanual interactions, I opted to use multitouch input.

The second reason for using multitouch input, which I discuss in chapter 1, is in response to the nature of LiquidText's document representation. As I outline above, the document representation is predicated on having many degrees-of-freedom in its visual, spatial presentation. Properties such as the positions of pieces content, scales of pieces content, and associations between content, may all be put under the user's control at a fine granularity. Manipulating properties such as these could be done serially. However, it would be inefficient and unnecessary to do so. With modern multitouch input, we can effectively expose a large number of properties for parallel control. Multitouch is also an ideally suited means of parallel input for LiquidText because it has a natural spatial mapping. In other words, with LiquidText, the degrees of freedom being controlled are, in many cases, visual/spatial attributes of content. The input coming from a multitouch system consists of points overlaid on a display, and as such has a natural mapping into the display space. So while a chorded keyboard certainly represents a form of parallelism, it generally does not have a natural mapping to a visual/spatial system, and would not be as appropriate for LiquidText.

4.2 Design Approach

Within the context of providing a high degree-of-freedom document representation, my goal in creating the specific design for LiquidText was to overcome some of the difficulties inherent in paper-like documents as discussed above. In particular, I sought to focus on areas where the lack of flexibility seemed to be the most problematic, including the visual arrangement of content (in terms of original material as well as annotations) and also navigation through content (including both original and user-created navigational structures). Consequently, my design process had two goals: 1) create a system that embodies this notion of flexibility, and 2) support additional active reading functionality as necessary to allow for an evaluation. That is, in order to enable the core ideas of LiquidText to be able to be evaluated with actual readers, various ancillary functions or design elements would be necessary—such as proper font rendering, a balanced color scheme, etc.

Alone though, these two criteria would suggest multitudes of functions that could be added to LiquidText. As such, I focused my designs on supporting the specific type of final evaluation I was planning. That is, the summative evaluation of LiquidText described below is a controlled lab study involving only individually-performed, time-restricted, self-contained active reading tasks. As such, functions for, e.g., collaboration would not be added to this prototype, even though a real world active reading tool would require them. Thus, the feature set was governed by a balance between: 1) breadth, to embody many attributes of the flexible document representation concept; 2) depth, to

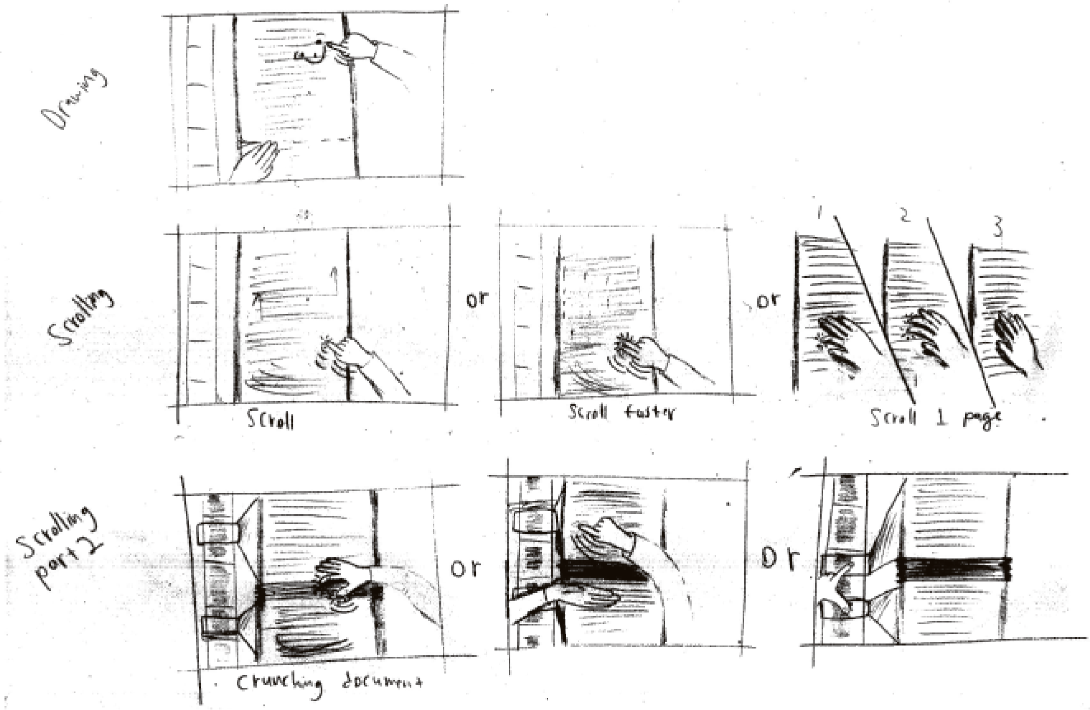


Figure 4.2. Design sketches that came out of the initial ideation corresponding to different LiquidText features, some of which were included in the final system.

make the prototype usable for limited, but real tasks; and 3) development simplicity, to facilitate short development cycles and an adequate number of design iterations.

4.2.1 Initial Design Ideation

I began the development of LiquidText by creating scenarios and accompanying requirements for active reading, grounded both in my own experiences as a reader as well as the prior literature. I also began exploring alternative metaphors for document representation—ones that would promote a user experience that felt fluid and flexible; not only allowing, but *inviting* users to step outside the bounds of predefined structure that computers can easily impose.

From a design perspective, to construct these metaphors, I found it helpful to explore forms and substances that exemplify malleability, such as putty or water. And while considering these ways of manipulating materials, I improvised interacting with imaginary systems inspired by such substances. Throughout this design process, I sought to brainstorm in the context of plausible scenarios, in order to lead my thinking toward designing a complete, integrated system, rather than just a collection of standalone interactions. Several resulting design sketches are shown in Figure 4.2.

Two guiding criteria for this phase of the design process were 1) to seek to include in my designs features that would support all of the major aspects of active reading, and 2), to explicitly focus on supporting those processes where paper-like representations fall short, as discussed above. And even from these initial designs, I sought to maintain a manageable scope by focusing on supporting the core processes of active reading. Thus I explored interaction with source texts and the creation of notes, as opposed to a larger active reading workflow including creating wholly new documents, sharing documents, and several of the other tasks discussed in the design workshops considered in the previous chapter.

4.2.2 Prototyping

After completing an initial system design, I sought feedback from other designers using a simple medium-fidelity prototype containing a limited set of interactions (like fisheye text zooming, and creating and grouping notes). This led to several semi-formal critiques where other HCI professionals offered feedback on my ideas and prototype. This provided feature suggestions and design criticism, and helped to inform the specific required feature set of the project.

In developing this initial prototype, I also chose to focus on single-document scenarios to keep the scope of the project manageable and the design goals tractable. Adler et al. likewise show this is not uncommon; of time spent reading/writing with multiple documents, about 40% had only one document being read (i.e., the other document(s) were writing surfaces) [Adler, Gujar et al. 1998].

4.2.3 Formative study

As explained in the previous chapter, my formative study consisted of an investigation into general issues of active reading, as well as a direct evaluation of the then-current LiquidText prototype. In this section, I describe the prototype evaluation in more detail.

In the context of evaluating LiquidText, the goal of this formative study was twofold. The first goal was to identify ways to make LiquidText a better, more complete embodiment of an active reading environment built around a high degree-of-freedom document representation. This potentially included identifying new functional requirements or otherwise making the system a better active reading system. The second goal of the study was to assess the system's basic usability properties, such as whether the user receives proper feedback in the UI, or can access the functions efficiently.

As noted in the previous chapter, the study included 20 participants, but in the context of evaluating LiquidText, two were used as pilot participants, so only 18 are considered to have completed the evaluation. The evaluation took place during the interview session, immediately after the general active reading interview. It consisted of the following phases:

- *Microsoft Surface Demo*: As discussed below, one part of this interview included participants creating their own touch-screen gestures for LiquidText functions. However, for participants who had no exposure to touch and multitouch interfaces, it could be difficult to envision what such gestures should look like. Indeed, [Wobbrock, Morris et al. 2009] describes participants creating gestures that would be unrecognizable to an actual touch-screen. As such, I sought to expose participants to a small collection of simple, common multitouch gestures. I chose to use gestures that are already becoming ubiquitous through consumer products like the Apple iPhone and Microsoft's Windows 7.

  In particular, I had demonstrated three applications to participants. The first simply showed what the Microsoft Surface cameras saw—proving that the system could detect only fingers or hands touching the screen and nothing happening further up in the air. The second application was a photo-sorter, allowing participants to move, rotate, and resize photos. The final application displayed a 3D model of a molecule; it allowed participants to scale, pan, and rotate the molecule in the plane, like the photo-sorter. However, it also demonstrated indirect multitouch interaction by providing a virtual trackball to rotate the molecule in 3D space. Thus, participants were shown basic direct and indirect multitouch interaction techniques, and given a sense of what types of gestures a multitouch screen can detect.

- *LiquidText Introduction*: In order to prepare participants for using and evaluating LiquidText, I showed them the basic visual layout of the application and explained what LiquidText is and what problem it is intended to solve. However, I did not show them the application in use, or describe specific features.
- *Gesture Development*: Here, I sought to determine the most intuitive gestures for the different LiquidText interactions. For this, I presented participants with a multitouch tablet which displayed an animation of a particular LiquidText interaction occurring (e.g., zooming the document). I then asked the participant to perform whatever gesture they would most expect to use to make that interaction occur. The tablet was instrumented to record the gestures performed.
- *LiquidText Tutoring*: In order to prepare participants to use LiquidText, I spent up to about 15 minutes tutoring them on using the then-current system. With one tablet PC in front of me and one in front of the participant, I demonstrated each interaction on my system and asked the participant to perform the same gesture on theirs. For ergonomics, each tablet was resting on an adjustable-angle stand on the table at which the participant and I sat.
- *Activity Preparation*: The actual active reading task participants performed was a critique: participants read a 4.5 page science article in LiquidText and wrote a critical

response using a traditional word processor. Naturally, LiquidText ran on the tablet PC's internal, multitouch display; the word processor ran on a second monitor connected to the tablet. Participants also had a wireless mouse and keyboard for interacting with the word processor (the keyboard would also be used to write comments in LiquidText).

In preparation for the active reading task, I gave each participant a gesture reference sheet showing several common LiquidText interactions with diagrams of the associated hand gestures. I also provided a detailed description of the critique task. Finally, I gave each participant about five minutes to read the directions, adjust their chair, adjust the angle of their tablet, arrange their workspace, play freely with LiquidText, etc. To encourage experimentation, I left the room during this interval.

To note, the tablet was a Dell Latitude XT2, with a 12.1", 1280x800 pixel display. The second monitor was a 1280x1024 pixel, vertically oriented, traditional LCD display. They ran Windows 7 and used WordPad for the word processor.

- *Active Reading Task*: When I returned, I loaded the designated article on the participant's tablet and instructed them to complete the task in 25 minutes. I sat several seats away from the participant but instructed them to ask me for help if the application malfunctioned, if they forgot a gesture, or if they experienced any other difficulty.
- *Closing Interview*: After the active reading task, I debriefed each participant for about 10 minutes. I began this semi-structured interview by letting participants freely describe their experience. I then asked specifically about a variety of topics, starting with the active reading task itself: was it too easy or hard, was it representative of their work, etc. I then enquired about various aspects of the functionality of the system, such as navigation, annotation, what additional features they wanted, etc. I also discussed the learnability and memorability of the gestures, the ergonomics of

the space, and the like. I closed with questions about affect, and whether/how a tool such as LiquidText would be useful in the real world.

4.2.4 Redesign

After completing the formative study, I used the results of the evaluation of LiquidText and the general investigation into active reading to substantially redesign the system. To make the required changes, I assembled a small design team who worked with me to explore different ways to satisfy the new user requirements I observed.

As explained in more detail below, one of the most important problems with LiquidText to emerge from the evaluation was the multitouch hand gestures themselves. While most of the gestures did not present problems for users, a small number were very challenging. As such, one of the key tasks facing the redesign was rebuilding the gesture vocabulary so all gestures would be at least performable and memorable. But unfortunately, multitouch input is still relatively new and LiquidText's gestures atypical. Thus, to properly refine the gesture set, I conducted a second formative study.

4.2.5 Gesture Refinement Study

After my design team and I developed a tentative new gesture set, I conducted a second formative study designed to evaluate and refine basic usability properties of LiquidText, particularly the gestures. While the first formative study was concerned with both usability and fundamental usefulness in the active reading process, this study was only concerned with the former. As such, its basic structure was to show users how to perform the various functions in LiquidText, and then ask them to perform those functions, observing where they struggled.

Also in contrast to the prior study, the objective here was not just to identify usability problems, but to try to resolve them. As such, I conducted many iterations, typically allowing only two or three participants to interact with a given LiquidText user interface before modifying it based on the challenges that were identified. Thus, rather than strive for the consistency required for high levels of scientific generalizability, this study sought to identify as many problems with as many variants of the LiquidText UI as feasible with the available resources.

The study included 13 participants from the Georgia Tech community, 3 of whom were female. Participants were compensated with $10 gift cards. For each of these participants, the study consisted of several phases, including:

- *Introduction*: To begin, I explained the purpose of the study and administered a demographic survey. I asked to participants to state any difficulties they have throughout the study, even if they can overcome them.
- *Tutoring*: Since the LiquidText prototype is only being used in a controlled, laboratory environment, I did not make a strong attempt to ensure that all functionality would be discoverable. As such, I spent approximately 20 minutes teaching participants the different functions in the system. I taught each function by demonstrating it on a multitouch tablet located in front of me. I then asked the participant to perform the same operation using a tablet located in front of them. As in the previous study, both tablets were Dell Latitude XT2's with 12.1" 1280x800 displays. Unlike in the previous study, the participant's tablet was flat on the table, rather than on an angle-adjustable stand.
- *Disruption task*: When participants were asked to recall how to perform the various gestures, I did not want them to rely on sensory or working memory to recall the training they just received, but rather rely more on long term memory. Thus, I sought to fill their working and sensory memory with unrelated material by giving the participants a multi-digit multiplication problem (a common technique for this type of situation, e.g., [Araya, Akrami et al. 2002]). Specifically, I asked them to multiply two 2-digit numbers, though they were allowed to use paper to find the solution.
- *Testing*: After completing the disruption task, I tested the participant on how well they could perform the functions, again admonishing them to state any difficulties

experienced. I then instructed the participant to perform each of the functions I had just taught them, but in a different order. I described the functions to the participant using terminology similar to what I used in the training A co-experimenter video-recorded and took notes.

- *Conclusion*: At the conclusion, each participant answered a questionnaire about the LiquidText functions. For each major function, the participant was asked whether it as easy to physically perform the gesture, easy to remember the gesture, and whether the gesture was a good match the function with which it was associated.

4.2.6 Final Redesign

While the Gesture Refinement Study provided the opportunity to make many small changes between iterations, a variety other problems were identified that did not need to be dealt with on so short a timetable and were deferred to a final design and development phase. These included three types of issues: 1) those where a solution was not in question—such as bugs, or simple design shortcomings, where additional evaluation cycles would be superfluous; 2) those which were too complex to implement in the time available, such as unlimited undo; and 3) functions that were not essential to active reading per se, but were necessary for conducting the summative evaluation. Of this latter category, perhaps the most important example was a simple word processor. As described below, the summative evaluation requires users to read documents with LiquidText and write responses. As such, I incorporated very simple word processor functionality into LiquidText. The resulting system is the final version of LiquidText which was used in the summative study and detailed in the next chapter.

4.2.7 Summative Study

The final, summative study is described in detail in Chapter 6, but here I provide brief synopsis. The goal of the summative evaluation was to answer my latter three research questions; that is, to investigate LiquidText's impact on the 1) subjective experience, 2) reading process, and 3) objective results of active reading. To investigate these, I conducted a controlled laboratory study with twelve participants. Similarly to the formative study, I had participants perform an active reading task (writing a critical response to an article) using LiquidText, but also had them perform a similar task using a control medium. I assessed LiquidText's impact through questionnaires and interviews, grading of the critical responses, and direct observation of participants' reading processes on the two media.

The larger goal of this study was to investigate LiquidText's general approach to representing documents, and this is discussed in Chapter 7 in detail. However the study also provided feedback on the specific interactions and design choices included in the final version of the system, and I report on this subset of the study's findings in the next chapter.

CHAPTER V
LIQUIDTEXT SYSTEM

While the previous chapter provided an outline of the methods used to develop LiquidText, this chapter discusses the actual system, and its evolution, in detail. To do this, I explain each of the features in LiquidText, describe the purpose it serves, and discuss its evolution through the different design phases. I also consider several features that have been removed from the system and explain why they were taken out. First though, this chapter begins with a brief overview of the final, completed LiquidText system.

5.1 Overview

The final version of LiquidText used in the summative evaluation is actually designed to run on a computer configured with two monitors. One monitor, typically the left, displays the LiquidText window (Figure 5.1), which is the actual multitouch document reading application used to evaluate my research questions. The second monitor displays a very simple word processor app (Figure 5.2) that can communicate with LiquidText.

Within the LiquidText window, the user is presented with two panes: The left contains the main document the user has loaded for reading (5.1 C); a range of interactions (described shortly) allow the document to be panned, scaled, highlighted, partially elided, and so on. The large *workspace* area (5.1 D) provides space for any comments or document excerpts the user may have created; here, such objects can be grouped and arranged, and the workspace itself can be panned or zoomed. Users can also create multitouch fisheye lenses to zoom only certain areas. Finally, users may also create navigation links among the various forms of content in the system, such as from the main document to comments or excerpts in the workspace.

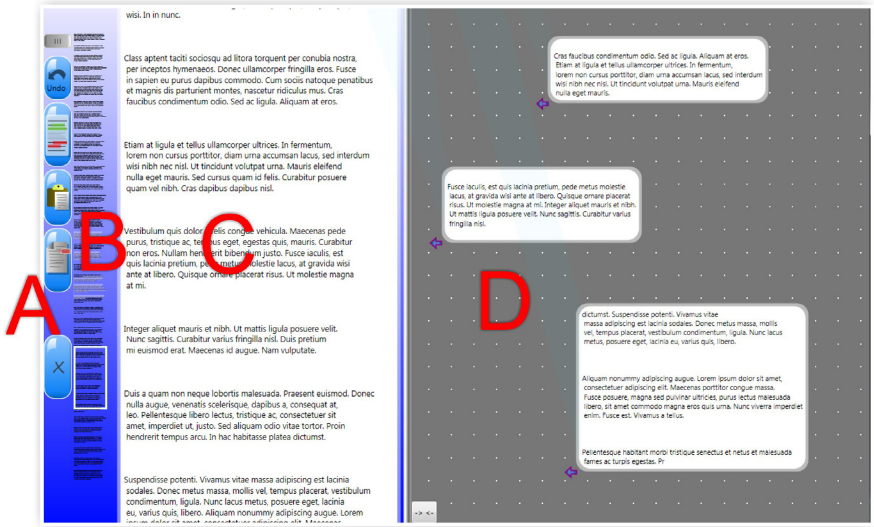


Figure 5.1. Overview of LiquidText screen: A) button bar, B) preview pane, C) main document, D) workspace, E) color palette, F) Dog-ear area.

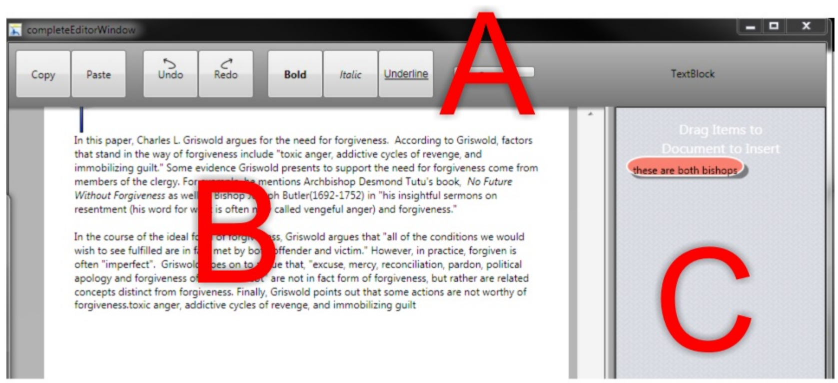


Figure 5.2. Overview of the word processor linked to LiquidText: A) Button bar, B) document being written by user, C) area where text selected in LiquidText appears.

Throughout the LiquidText user interface, some basic interactions reuse a number of common gestures and concepts that appear in other touch applications (e.g., [Wu and Balakrishnan 2003]). For example, the user can position objects, such as comments, excerpts, and documents, just by dragging with a fingertip. The user can rescale an object by using two fingers to pinch or stretch it, or she may rotate it by using two fingers to twist it. The preview pane next to the document (Figure 5.2, B) provides a scaled-down view of the document, and the user may simply touch it to navigate the document to the corresponding point. Additionally, LiquidText provides unlimited undo, and allows the user to hold down a delete button (last button in Figure 5.1 A) while touching objects to delete them.

The word processor app (Figure 5.2) is not fundamentally part of LiquidText itself, but was included to support the summative evaluation of the system. As described in detail in Chapter 6, the presence of the word processor allows the user to write a response to the document being read in LiquidText, without having to switch to a traditional, mouse-based writing application. This word processor, however, is simple, containing a button bar (5.2 A) for performing only basic font changes, undo/redo, and copy/paste. And in addition to the document writing space (5.2 B), it contains an area where text can be easily brought over from the LiquidText window. And since this word processor is meant to be used alongside LiquidText itself, it supports several basic touch gestures, such as text selection and scrolling, again, so that no mouse is required.

In the following sections I discuss the details, rationale and evolution for the central aspects of the design of LiquidText, as well as a brief explanation for the design of the word processor. And while discussing certain aspects of the LiquidText design, I also summarize select, relevant findings from the summative study to provide a snapshot of how these features were ultimately used in practice (note though that the complete findings of the summative study are presented in Chapter 7).

The remainder of this chapter begins with a discussion of LiquidText's design organized according to the major active reading processes described in Chapter 2: content layout and navigation, content extraction, and annotation. This is followed by a discussion of the word processor, several functions that were not included in the final system, implementation, and design implications.

5.2 Content Layout and Navigation

LiquidText provides a number of novel features intended to support the active reading processes of content layout and navigation. These features allow users to access an existing text *linearly* in its original form (as shown in Figure 5.1 C), yet leverage a range of interactions to selectively view multiple regions of text, and to move quickly to different parts of a text. The major LiquidText features for content layout and navigation are Collapsing, and Fisheye Views in the workspace area.

5.2.1 Collapsing

Among the most important aspects of layout in active reading is the need for parallelism, such as viewing multiple pieces of a document at once [O'Hara 1996; Morris, Brush et al. 2007]. But as intuition suggested, and my formative study revealed, viewing disparate areas of a document in parallel is often difficult, requiring frequent flipping back and forth. To better support this in LiquidText, I was motivated by elastic substances which can be selectively compressed or expanded—suggesting a document that lets one compress or shrink some areas to bring text from disparate areas of a document together—resulting in a sort of 1-dimensional fisheye. The visualization is thus similar to [Hornbaek and Frokjaer 2003], and the interaction similar to [Zeleznik, Bragdon et al. 2010], but in a very different context.

The resulting visualization, called "collapsing," hides select regions of the document by reducing selected rows of text to horizontal lines; as a result, disparate areas of the text may be brought into proximity. Since text rows are not removed completely but only vertically reduced, the user is given a cue indicating the amount of text hidden. Additionally, multiple regions of text may be collapsed at once, letting the user choose precisely which portions of the document are visible in parallel. Interactions involving this effect are used throughout LiquidText.

Since this collapsing process is applied to a vertical, linear document (Figure 5.1 C), LiquidText uses a three finger vertical pinch gesture to provide a natural mapping for this interaction (Figure 5.3). This simple gesture offers several degrees of freedom for controlling which text is elided. First, the user can control whether text above or below their fingers is collapsed, by either pulling the top of the document down, or by pulling the bottom of the document up. Likewise, the user can control the amount of text that is

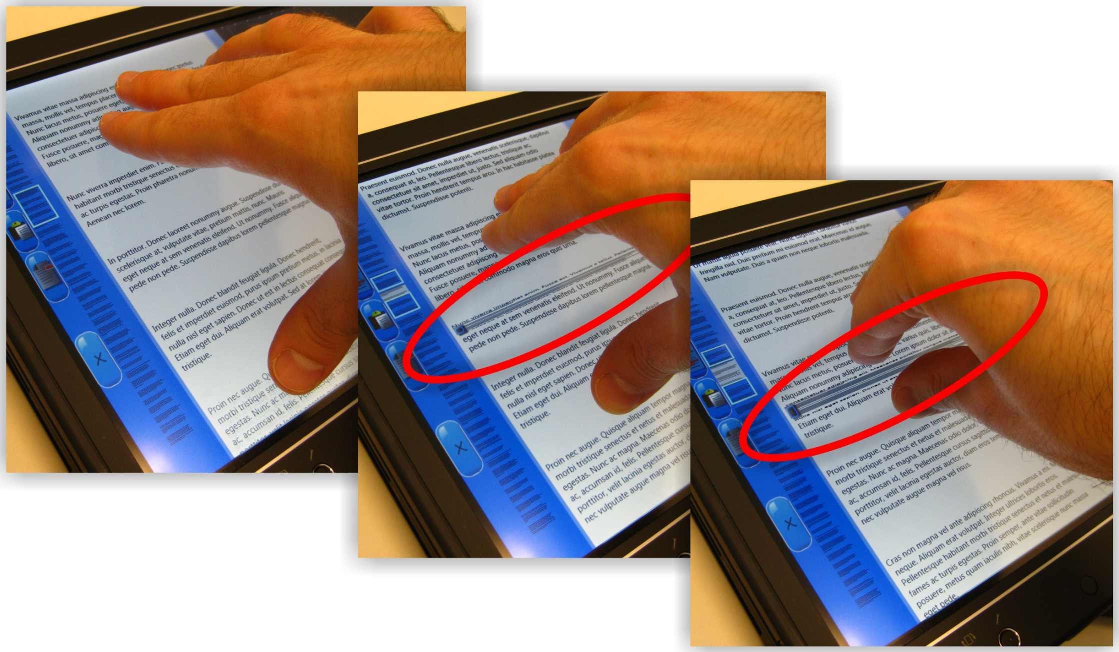


Figure 5.3. Three stages of an increasing amount of text being collapsed together.

collapsed based on the distance through which their fingers are moved. As one would expect, the user can move their fingers apart to un-collapse text that has already been collapsed.

The preview pane (Figure 5.1 B) plays several roles in collapsing as well. First, as described below, it can be used to directly collapse text. Second, touching the preview pane instantly eliminates all collapses (in addition, of course, to instantly scrolling the document to the selected location). This gives the user a fast way to restore their document if they made many collapsed areas. Third, the preview pane renders collapsed rows of text with a gray background (Figure 5.3), as an indication to the user that the given text is still present in the document, but simply not visible.

The overall collapse interaction, however, raises the design question of what should happen when the user scrolls a partially collapsed document. For example, if the user scrolls text above a collapsed area, either 1) the entire document could scroll, including the collapsed area. Or 2), only the text above the collapsed area could scroll, adding text to, or removing it from, the collapsed area. Under the first option, the collapse is *a part* of the document itself, and can be used to conceal irrelevant material. Under the second option, the reader effectively has two windows into the document, each of which can be scrolled independently. Both choices have advantages, but I selected the first option as it better fit the metaphor of pinching/stretching the document.

By contrast, the latter option is similar to systems like Adobe Reader, which show disparate regions of a text by allowing the user to create multiple, independent views into the document via a split pane. But that approach to multiple, independent views comes at a cost, offering little clue as to the relative order and distance between the document regions; this disruption to the document's underlying linearity may also interfere with a user's sense of orientation [O'Hara 1996]. Nonetheless, in either approach, LiquidText's collapse interaction guarantees that visible regions are always in the order of the original document, and provides a visual cue as to how much text has been collapsed between two visible regions as well.

While this basic collapsing interaction provides a useful way to hide irrelevant text, and bring disparate document regions into proximity, manually pinching a long document together to view, say, the first and last lines of a book, is tedious. LiquidText thus provides two indirect ways to collapse a text: first, touching the preview pane with

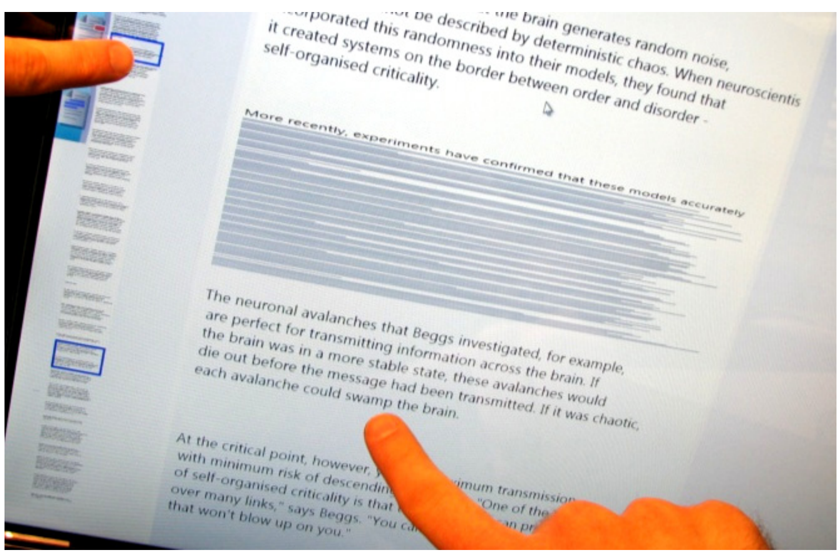


Figure 5.4. Holding one finger on document while scrolling another finger on preview pane collapses document to show both locations.

two or more fingers causes the document to automatically collapse to show all the indicated areas at once. Alternately, holding down a desired part of the document with a first finger—effectively holding it in place—while using a second to touch the preview pane, causes the document to collapse so as to keep the area under the first finger visible, while also showing the area indicated by the second finger (Figure 5.4). And in contrast to traditional document viewing software, in which users must create separate panes and scroll them individually, this functionality lets a user view two or more document areas with just one action, parallelizing an otherwise serial task.

Evolution

The collapse interaction itself was one of the first functions in LiquidText, and was thus available to evaluate in the formative studies. The first formative study revealed little difficulty with the interaction, but did show that users tended to see a dichotomy between using collapse and creating excerpts (described below), since they could be used for similar purposes. Notably though, two participants used collapsing as a way to conceal irrelevant material, even though I only presented it as a way to bring disparate material together—suggesting the decision to make collapses part of the documents themselves was of value to users.

The *gesture* used for collapse though, underwent more revision. In the formative study, the collapse gesture required only two, rather than three fingers, and participants found this very natural. But to distinguish this collapse gesture from the rescale gesture (also a two-finger pinch, used to resize the document, comments, or excerpts), the latter had to be performed horizontally, while collapse was performed vertically. This posed problems for formative study participants. Partly, this was because the pinch-rescale gesture used on most touch screen devices is orientation agnostic, and participants had difficulty recalling that LiquidText behaved differently. But the greater problem was the ergonomics of the horizontal pinch motion. As such, my design team and I redesigned the gestures such that a two finger pinch always maps to rescale, and a three finger pinch always maps to collapse. As can be seen in Figure 5.5, zooming individual objects did become physically easier for participants in the later iterations of the gesture refinement study, and our observations of participants difficulties suggested the gesture change did work. Positively, the change to the collapse gesture generally did not pose problems either.

The final change made to collapse was to change the row-scaling factor. Initially, all rows were scaled down to 1-pixel in height, but when large amounts of text were collapsed, this caused the collapsed area to consume a large amount of the display space. After a study participant expressed concern about this, I changed it such that the total height of a collapsed area increases with the log of the number of lines of collapsed text.

Summative Study Feedback

In the summative evaluation, feedback on the collapse functionality was generally positive, but also fairly sparse. Only one participant provided an assessment of the collapse functions in the interviews, referring to it as “amazing.” But while collapse received little criticism, and is one of the more visually distinctive features of the system, it was only used by two study participants. Participants explained that for the main purposes collapse is meant to serve—comparison and concealing irrelevant content—they chose instead to create excerpts, which are described below. However the reason for this preference was not made clear through the interviews.

5.2.2 Fisheye workspace.

As noted, the workspace area to the right of the document allows users to freely arrange comments and excerpts (described shortly). These can be rescaled to make them larger and more readable, or smaller so they take up less space. For the version of LiquidText used in the formative study, I expected that, if users ran out of space, they

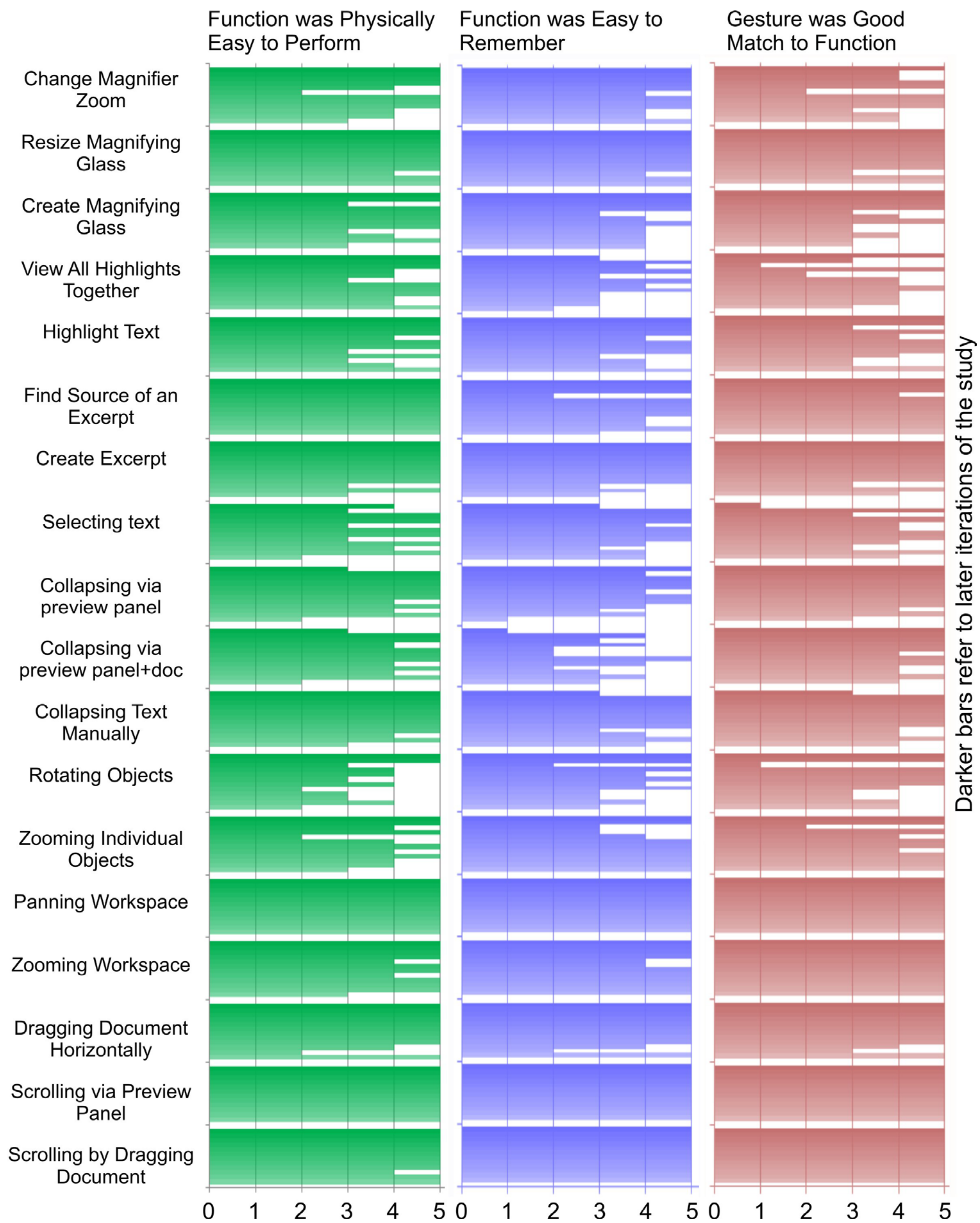


Figure 5.5. User reactions to the various LiquidText gestures over successive iterations of the Gesture Refinement Study

would simply rescale individual items or groups of items so they could all fit in the space available.

During the formative study however, I found this was not the case. First, users commented on the dearth of workspace available to them on the 12.1" tablet which ran the prototype. Likewise, the participatory design workshops later in the formative study indicated that users required a large space in which to see all their documents at once—

and to get an overview of any cross-document relationships. Scaling individual groups or objects was clearly not adequate.

To overcome these challenges, I considered several alternatives for letting users work with larger numbers of objects effectively. Obviously, a physically larger display would be helpful, and so I replaced the 12.1" display with a 17" display, as described below. But even 17", the largest display I could practically implement, it is still a fairly small working space. Additionally, just using arbitrarily larger displays is not a good general solution outside the bounds of this study, since larger displays trade off against portability (something formative study participants strongly required, as they reported often shifting between various public and private working areas). As such, a visualization solution was required to complement the larger display.

For this, my design team and I considered providing some form of organizational structures, such as allowing users to make hierarchical outlines of comments and excerpts in which tiers can be shown or hidden as needed. Outlines have certain downsides, however: they impose a strict organizational structure that may be inflexible, and they also privilege hierarchical over spatial relationships.

Instead, for the final system, I settled on the notion of a quasi-infinite, continuous workspace for comments and excerpts, extending beyond the display. This workspace can be panned and zoomed, thus supporting spatial overviews of comments and excerpts, and maintaining consistent spatial relationships among the objects within the space. Since simultaneous viewing of multiple pieces of content is important in active reading, I considered a number of approaches to supporting this functionality in the workspace region. One example of this was to allow regions of the workspace to be "collapsed" in the same way that LiquidText allows for documents. This could be used to bring distant areas of the workspace into proximity, and rely on an interaction and metaphor already present elsewhere in the UI. But while this idea is appealing, the technique was most suited to one-dimensional spaces. Collapsing in a large 2D space would hide a great deal

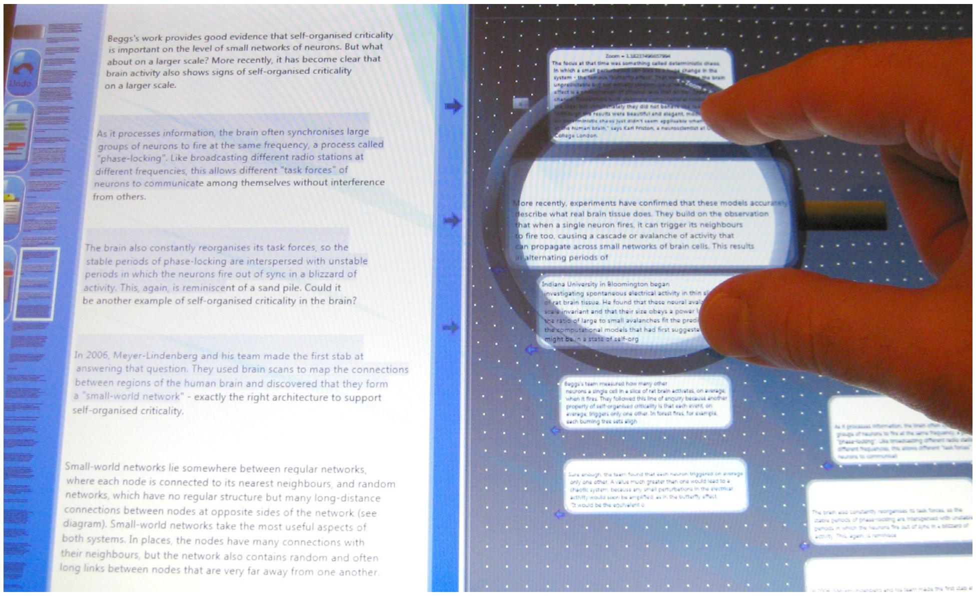


Figure 5.6. Multitouch fisheye lens function.

of content unnecessarily and, if vertical and horizontal collapses were used at once, could easily become confusing.

To avoid these downsides, I instead developed a novel type of fisheye effect that would allow the user to see the overall context of their workspace while zooming in on one particular area of interest. To do this, the user can create one or more fisheye "lenses" in their workspace (Figure 5.6). The lens applies a fisheye distortion to each of the objects in the workspace; enlarging those within the circumference of the lens, and shrinking the surrounding space so as to reduce object overlap—thus generally allowing all objects to remain visible. The distortion is consistent within each object so text remains readable as well. The disadvantage of this intra-consistency, however, is that it makes it harder for the user to develop a mental model of the fisheye, since they only see its effect at a very small number of points at any given moment. For a simple magnifying glass, this might be acceptable, since people use those frequently in the real world. But a fisheye lens is less common, involving some areas being enlarged and others shrunken. Thus, I included a grid of dots in the background of the workspace which are also affected by the distortion. Seeing the effect of the distortion at so many points is therefore intended to give the user a more complete sense of what the fisheye lens is doing.

In the final version of the system, the workspace is panned or zoomed using typical drag/pinch gestures. Fisheye lenses are created by tapping with three fingers; once created, they can be moved by dragging and resized by pinching or stretching the lens border. The magnification level is controlled by rotating the lens—akin to the zoom lens of a camera. Thus, the user can control all the essential properties of the distortion in tandem and, using multiple hands, for two lenses at once. Thus, while fisheye distortions have been deeply explored, this approach combines two unique features: consistent scaling within each object, and the ability to create and manipulate the distortion's attributes in parallel through multitouch.

Evolution

While the Fisheye Workspace was not tested in the formative study, since it only came as a result of that study's findings, the gesture refinement study did reveal several problems with the magnifying glasses' interactions.

One of the challenges was balancing unobtrusiveness with usability. If the magnifying glasses were allowed to reside beneath the content they were magnifying, they could easily be lost underneath that content. But if they were kept above the content, the magnifying glasses' borders would themselves occlude some of the content being read. My initial solution was to keep a magnifying glass's handle above other content, but let the rest of the object settle below—thus giving the user a "life line" to retrieve the magnifying glass. Unfortunately this did not work well with users, who visibly struggled to keep the lens elevated by holding its handle while using another finger to rotate or resize the lens. For the final design, I used a different approach whereby the magnifying glass is always above the content, but simply becomes less occlusive when not being adjusted. Applying the principles of [Baudisch and Gutwin 2004], any portion of the magnifying glass that overlaps another object is displayed as translucent and blurry when not being touched, so that the user can still read any underlying text. When the user

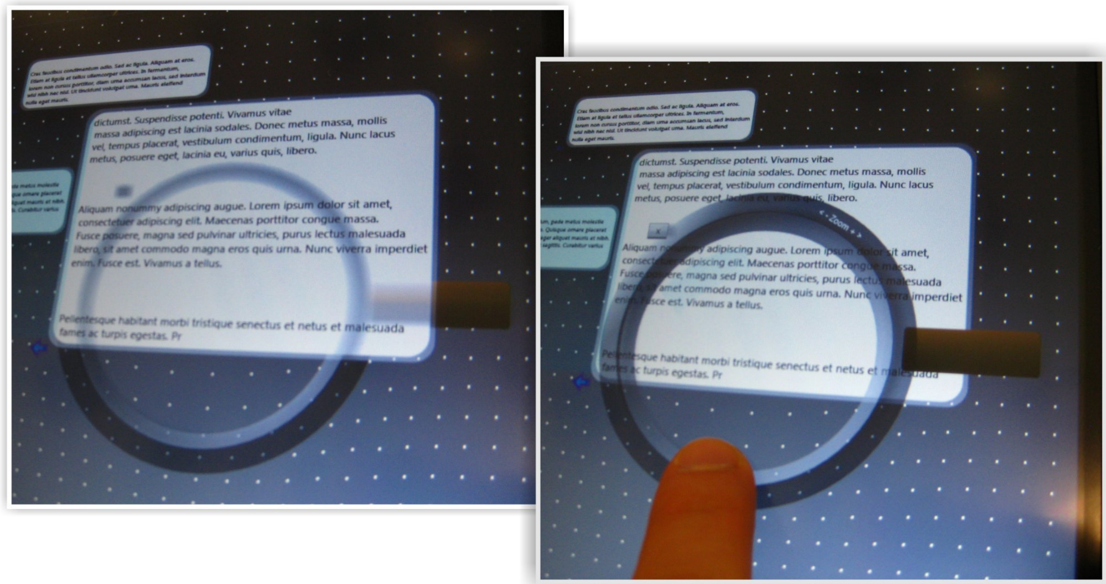


Figure 5.7. Magnifying glasses are shown blurry and translucent where they overlap other objects, but they become solid while the user touches them.

touches it, it becomes sharp and more opaque so the edges can be clearly identified (Figure 5.7).

A variety of simpler problems with the magnifying glasses were also resolved through the gesture refinement study. These included situations where the user could accidentally create several magnifying glasses in the same place and, more seriously, situations where users felt disoriented because of the intensity of the fisheye distortion. To resolve the latter, my first change was to minimize the times when the user has to adjust the magnification, since that often increased the disorientation; I did this by choosing reasonable default magnifications based on the overall zoom level of the workspace. Second, I limited the zoom level of each magnifying glass to a more conservative value. As evident from both user comments and the questionnaires (as shown in Figure 5.5), the final gestures worked adequately well. Nonetheless, participants did continue to experience some difficulty maintaining a clear, oriented understanding of the distorted environment (especially as they moved objects through the distorted space), and so it may be helpful to explore refining the distortion itself in future research.

Summative Study Feedback

In the summative study, the workspace received a strongly positive reaction. As I explain in Chapter 7, participants used the workspace as a central part of their active reading process, employing it as a repository for their excerpted content and their comments. In the interviews, participants praised the size of the workspace, and described it as helping them organize their thoughts. But more prominently, participants felt the workspace added to their awareness of their materials, since it let them bring all of their materials, such as excerpts and comments, into one location.

But while the workspace played an important role in participants' reading and received positive feedback, some of the functions available within the workspace did not. While eight[4] participants used the functions for panning the workspace, only four[4] used the workspace zoom functions, and just two[4] used the magnifying glasses. In the interviews, participants explained that part of the reason for this was that the available screen space was sometimes adequate on its own. Two participants, for example, noted that they might have used the more of the workspace navigation functions had the article or the task been longer. A third likewise noted that panning was adequate and zooming was not necessary.

Another reason some workspace functions, particularly the magnifying glasses, may not have received more use was because some users struggled with them. Two described the magnifying glasses as difficult to rotate for changing the zoom level, which was consistent with me own observations while training participants to use LiquidText, in that they often seemed to have difficulty controlling the lenses effectively. There are also hints that the fisheye distortion itself may have been jarring at times, as one participant wanted to use the magnifying glasses but felt the fisheye effect, "distorts [the workspace]

[4] Due to a data loss, this number is based on data gathered from only 11 of the 12 users.

so much that it's hard to use." Still, most participants did not indicate this, so it is unclear whether this was generally the case.

5.3 Content Extraction

As one of the central processes of active reading, extracting textual excerpts from documents serves to support aggregation and integration of content from across one's documents. It lets the reader maintain peripheral awareness of key material, and explore alternate organizations of an author's content [O'Hara 1996; O'Hara, Smith et al. 1998; Marshall, Price et al. 1999]. But, in most present approaches to active reading, content extraction has required considerable effort: paper-based active reading may require copying, scanning, or rewriting portions of content in a notebook; even computer-based reading may require one to open a new document, copy and paste text into it, and save it. With LiquidText, the goal was to create a fast, tightly-integrated set of mechanisms for content extraction that would not distract from or interfere with the active reading process itself [O'Hara 1996].

To devise an intuitively appealing, flexible interaction to support content extraction I again sought to draw on physical metaphors. Imagining a document as puttylike, I approached extracting content as analogous to *pulling it off* of the document. The two major parts of this interaction are, first, selecting the text to be excerpted, and, second, actually transforming that selected text into an excerpt.

5.3.1 Text Selection

Creating an excerpt begins with selecting the desired text. In the final version of LiquidText, I selected a simple tap-and-drag gesture that is also used in some iOS apps. Thus, the user begins the selection process by putting down a finger at the start of the intended selection, lifts her finger and then lowers it again, then drags her finger to the desired endpoint.

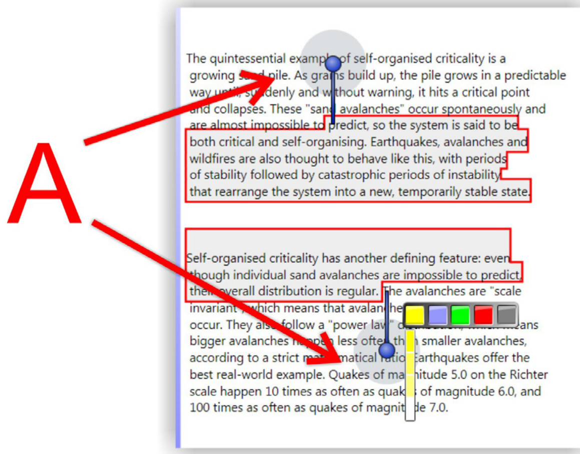


Figure 5.8.

Since selection can sometimes be difficult, LiquidText includes several simple features to help the user make an accurate selection. The first is to give the user prompt feedback as they start a selection, which is done by displaying a cursor where they initially put down their finger. Consequently, if the selection is starting in a location other than what the user expected, the user can pause for a moment and perform the gesture over again. Additionally, since it can be difficult to raise and lower a finger quickly and with precision, the start point of the selection is where the user initially puts her finger down, rather than where she puts it down after the tap. To further help the user select the start point correctly, since users rarely want to begin a selection in the middle of a word (since they cannot edit), LiquidText automatically moves the starting point of the selection back to the nearest word boundary. If the user completes their selection, but finds it to be incorrect, there are selection modification handles (Figure 5.8), which are standard on iOS apps, shown around the selected text. And finally, if the selection is completely a mistake, the user may double-tap outside the selection to remove it and any other selections in the document.

Evolution

In the version of LiquidText used in the formative study, there were two ways to select text: 1) an easy but modal approach in which the user held down a soft button with one hand while dragging a finger over the desired text, and 2) a non-modal approach, where the user puts a thumb and forefinger together just below the start point of the selection, and moves those fingers to the endpoint. This way, the user did not have to shift their gaze away from the document to invoke a selection mode.

Unfortunately, the formative study revealed disadvantages to both approaches. Several users disliked the bimanual requirement of the modal approach, and the non-modal approach was difficult to perform reliably. This inconsistency was due in large part to the touch sensors, which would sometimes confuse two closely spaced fingers for one large object. Interestingly, this problem appeared for certain users but not others, even if they performed the gesture with comparable finger spacing.

Rather, the gesture that users wanted for text selection was to simply drag a finger over the text—the same gesture they preferred to use to move objects and scroll the document. Thus, I sought a gesture distinguishable from, but very similar to, simple dragging. One option was to distinguish the gesture through time—to have the user put a finger down at the start of the selection and pause until the system enters selection mode. This is similar to what is done in iOS, but introduces an awkward delay into one of the most important interactions in the system (since so many other interactions depend upon it). Ultimately, I chose the tap-and-drag gesture, since it seemed indeed to be very much like dragging, but with only one additional movement at the beginning.

While users generally responded far more positively to this gesture than the earlier selection interactions, problems persisted. As can be seen in Figure 5.5, users struggled with physically performing the gesture throughout the gesture refinement study. The largest area of difficulty was in accuracy—being able to reliably select the desired starting point. My design team and I explored several alternative gestures, but ultimately found nothing that would be clearly better and would not conflict badly with other important LiquidText functions. Instead, we took the approach of keeping the current selection gesture while making it easier to recover from selection errors—hence adding the selection cursor and handles described above.

Summative Study Feedback

In spite of the iterations of testing, of the various functions in LiquidText, the summative study found that text selection was the most problematic. In the interviews, five of the twelve participants pointed to text selection as difficult (one claimed that it became easier with practice, though), while a sixth criticized selection in the questionnaires. Three of these stated that text selection was the hardest part of the using LiquidText.

The reasons selection was challenging were often attributed to lack of precision—participants struggled to set the selection boundaries at the specific word or line that was intended. While the study did not reveal a clear solution to this problem, several participants did agree that using a stylus for selection—even if still using touch for other tasks—would have been advantageous. While sensing both pen and touch input was technically infeasible for this research, current touch sensing technologies would make this a viable research direction for future work [Hinckley, Yatani et al.].

5.3.2 Excerpt creation

Once a span of text has been selected, users may create a duplicate of it known as an excerpt. Following the putty metaphor, the user creates an excerpt simply by dragging the selection with one finger into the workspace until it snaps off of the document (Figure 5.9). The original content remains in the document, although it is tinted slightly to indicate that an excerpt has been made of it.

The creation of a selection begins when the user starts to drag their selected text (Figure 5.9 A). At this point, an excerpt is created, but it is shown translucent and with a dotted border. This indicates that the excerpt is “tentative” and will disappear if the user releases it (Figure 5.9 B). This is to reduce the penalty if the user accidentally brushed a

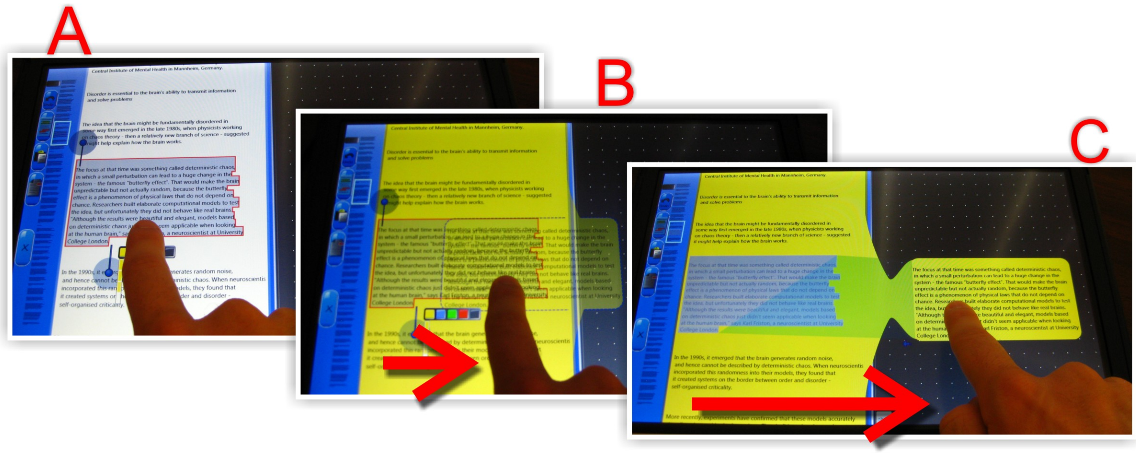


Figure 5.9. A) Beginning to drag a selection to create an excerpt. B) Dragging selection as indicated by arrow to create excerpt—excerpt is still considered "tentative" and so transparent. C) Excerpt has been dragged far enough to no longer be "tentative."

finger over the selected text, or dragged over it with a different intention (e.g., to drag the document). If the user continues to drag the excerpt, it will become opaque, indicating that the user can safely remove their finger once it is positioned to their liking (Figure 5.9 C).

As seen in Figure 5.9 C, the excerpt is initially attached to the original document. If left attached, the excerpt will be treated as a child object of the document, such that moving the document will move the excerpt as well. If the user continues to drag the excerpt away from the document, the putty-like connection between them will stretch until it snaps, and the excerpt is treated as an independent object.

Evolution

Excerpt creation was generally a well-received function throughout the formative study and the gesture refinement study. Participants used it extensively in the formative study, often in place of collapsing, since both can be used to compare two pieces of text side-by-side. Nonetheless, users expressed some challenges with the original gesture, which required holding the document itself in place with one finger, while using another finger to drag the selected text to create the excerpt (at which point the first finger could be lifted). The original motivation for this more complex gesture was to allow the user to scroll the document even by dragging text that had been selected, and to keep the excerpt creation gesture from conflicting with an earlier gesture used for highlighting text (which involved brushing a finger over selected text). Also, in the spirit of keeping the cost of recovering from an error comparable to the ease with which it is made, I was afraid that single finger dragging would allow small brushes or slips of a finger to create erroneous excerpts littering the user's documents.

Nonetheless, although users could learn the more complex gesture, it was a substantial burden, and so I switched to the single-finger drag described above. The purpose of the "tentative" excerpt state, then, is to try to raise the bar to erroneously making an excerpt, by forcing the user to drag it a significant distance before it will persist. As can be seen from Figure 5.5, users were very satisfied with the excerpt creation gesture after it was fully refined.

5.3.3 Excerpt Manipulation

After creating excerpts, users must be able to organize and review them [O'Hara 1996; O'Hara, Smith et al. 1998]. To support this and my general design requirement of flexible content arrangement, excerpts can be freely laid out in the workspace area, twisted for rotation, and pinched or stretched for resizing. They can also be attached to one another (or to documents) to form *groups*. This allows users to create whatever structure they may desire with their excerpts, rather than the system imposing explicit relationships that may not be necessary or may not fit with users' models of the structure that they require.

In designing the ultimate appearance and behavior of these groups of excerpts, though, there was a tension between the advantages of structure versus flexibility. In one approach I considered, grouped excerpts move into a container-object, where they are aligned and sequentially listed, forming a sort of outline. This approach is visually simple

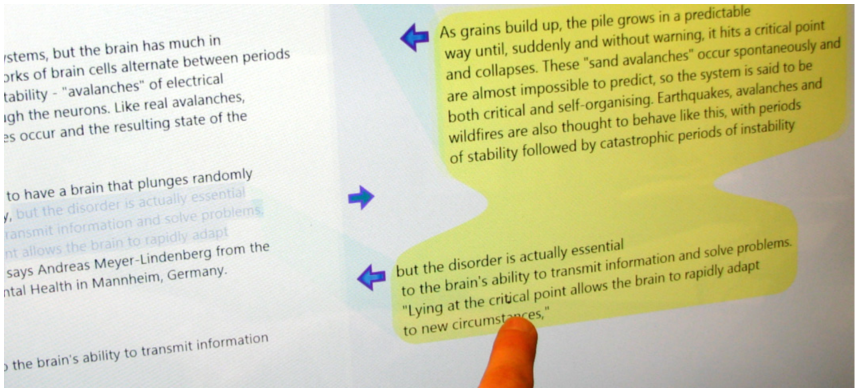


Figure 5.10. Attaching two excerpts to form a group.

and organized, especially for a small screen. Alternately, grouped objects could be positioned *arbitrarily* and freely within a group, and have visual links rendered between them to show they are grouped. This option gives users more means of expression, letting them indicate differences in group constituents by position/size/angle, but would likely be messier. I concluded however, that the workspace would adequately accommodate any potential disorganization, and so chose this latter option, informing the interaction's behavior with the putty metaphor used in excerpt creation.

Excerpts are grouped simply by dragging them together—which creates a fluid-like border surrounding all objects in the group (Figure 5.10). Pulling them apart stretches the border until it eventually snaps, separating the group. This choice of visual language was intentional: as discussed in the previous chapter, LiquidText was meant not only to allow for flexible interactions but to suggest to the user that the interface as a whole invited and supported them. Considering works such as [Ziemkiewicz and Kosara], which identified significant effects of design language on the semantics of how data are interpreted, I sought to provide visual design in LiquidText that would communicate a message of fluidity.

Evolution

Throughout the formative study, most of the functions for manipulating excerpts were well received, except for resizing and rotating (which likewise applied to other objects, like comments, which are described below). The version of LiquidText used in the formative study relied on two-finger *horizontal* pinching for resizing, and *three*-finger twisting for rotation. Both of these caused problems for users though, as revealed in the gesture refinement study (and shown in Figure 5.5). Thus, my design team and I ultimately decided to merge these two gestures into pinching or twisting with two fingers for resizing or rotating, respectively. And as explained above, I changed the gesture for collapsing in order to accommodate this change. As seen in Figure 5.5, the later gestures were indeed better received by users.

5.3.4 Excerpt Context

The final aspect of excerpts concerns the ways in which they relate to the text from which they were copied. The first relationship is their use as a means of returning to the excerpted text's original context in the source document—which is known to be vital for active reading [O'Hara, Smith et al. 1998]. In designing LiquidText, I explored two alternatives to providing such context: the first was *in-place expansion*, where the original context of an excerpt was made to appear around the excerpt, showing the context without disturbing the original document. The second was *linking*, where users can interact with an excerpt to cause the source document to position itself so the excerpt's context can be seen.

Each has advantages: in-place expansion lets the user keep her gaze in the vicinity of the excerpt and avoids disturbing the position of the original document. But if the user needs to do more than just glance at the source context, she will likely need to navigate the original document to that point anyway—which linking does immediately. Likewise, an in-place expansion on even a 17" display may well cover much of the user's workspace, raising visibility problems, especially if one wanted to see the contexts of two excerpts at once.

For this and other reasons I ultimately chose linking, which I implemented bi-directionally, so excerpts could link to sources and vice versa. Arrow buttons (shown in Figure 5.10) appear near both excerpts in the workspace as well as areas of the original source document from which the excerpts were taken. By touching the arrow button near an excerpt, the source document immediately moves to a new position in which it can show the excerpted text in its original context; likewise, touching the arrow button near the source document will move the excerpt object in the workspace into view. This mechanism effectively provides a way for users' own excerpts to be repurposed as a way to quickly navigate through the original source document.

An important aspect of the arrow buttons is that they take advantage of multitouch to support parallel use with LiquidText's other navigational interactions. First, multiple arrow buttons can be held down at once to view multiple contexts in parallel—this causes the source document to automatically scroll and collapse text as necessary show the selected areas. Likewise, the user can touch an arrow button on an excerpt while holding part of the original document in place, or while touching part of the preview pane, and LiquidText will scroll and collapse the document to show all indicated areas at once. This is similarly true for the excerpts themselves—if the user touches multiple right-facing arrows on the document in order to view several excerpts at once, the workspace will zoom out as required to make all of them visible.

5.3.5 User Feedback on Excerpts

Beyond the specific gestures involved in making excerpts, even the formative study participants showed generally positive reactions to the excerpting functionality as a whole. The idea of extracting content via touch was in-line with user expectations, as even before seeing the prototype system, some users expressed an interest in using touch to pull content out of a document, and several described the existing copy/paste process as being laborious. Perhaps as a result, of the various functions in the prototype, the ability to create excerpts received the most strongly positive feedback, with eight formative study participants noting it as a feature they liked. In the formative study's active reading task itself, participants typically used excerpts as a way to aggregate content or to provide peripheral awareness. They noted that they pulled out quotes and even whole paragraphs so they could refer back to them later. And partly due to the support for grouping, several formative study users discussed its value in allowing them to reorganize the content of the document. One user described this as, "What [the prototype] was really helpful for was because I had kind of written my own version of the story on the side." Excerpting, therefore, seems to have been used in much the way it was intended, but its frequency of use allowed it to easily consume available display space—providing part of my motivation for adding the fisheye workspace (above).

In the summative study, with the refined gestures and workspace, excerpts appeared to take on an even more important role. On average, participants used excerpts more than any other annotation or note taking device—such as comments or highlights. Generally, as discussed in Chapter 7, participants used excerpts to create in the workspace a condensed version of the article they were reading. As such, excerpts were described as letting participants view only the most important areas of the document while eliding the rest, and as letting them make comparisons between what would otherwise be distant document areas. Consistent with the apparent importance of excerpts in participants' reading processes, all participants created at least one excerpt, and on average created 3.6.

In discussing excerpts, summative study participants noted several positive aspects of the feature. One aspect cited by three participants was retrieval, in that excerpts made it easier to re-find important pieces of information. As one participant explained, she likes to write out excerpts even with paper, but tends to lose them; with LiquidText they do not get lost since she has mechanisms (like zooming out) to find them again. Two participants also cited arrangement, in that excerpts could be freely organized and grouped in the workspace. Additionally, individual participants noted a variety of other factors, including the pleasantness of the excerpt creation gesture, the connectedness of excerpts to their sources, and the integration that comes from having the excerpts' source documents and the workspace in the same program. Thus, while individual participant reactions varied, excerpts received a strongly positive reaction from participants.

5.4 Annotation

Annotation is one of the most important, as well as abundantly investigated, areas of active reading. LiquidText supports two of the more extensively studied forms of annotation: comments and highlighting [Renear, DeRose et al. 1999]. It also supports a less common operation, direct linking, which I also consider to be a form of annotation. I discuss these three areas of functionality in turn.

5.4.1 Comments

As discussed in chapters 1 and 3, one shortcoming of paper is the constraint it places on textual annotations such as comments. Comments on paper must generally be

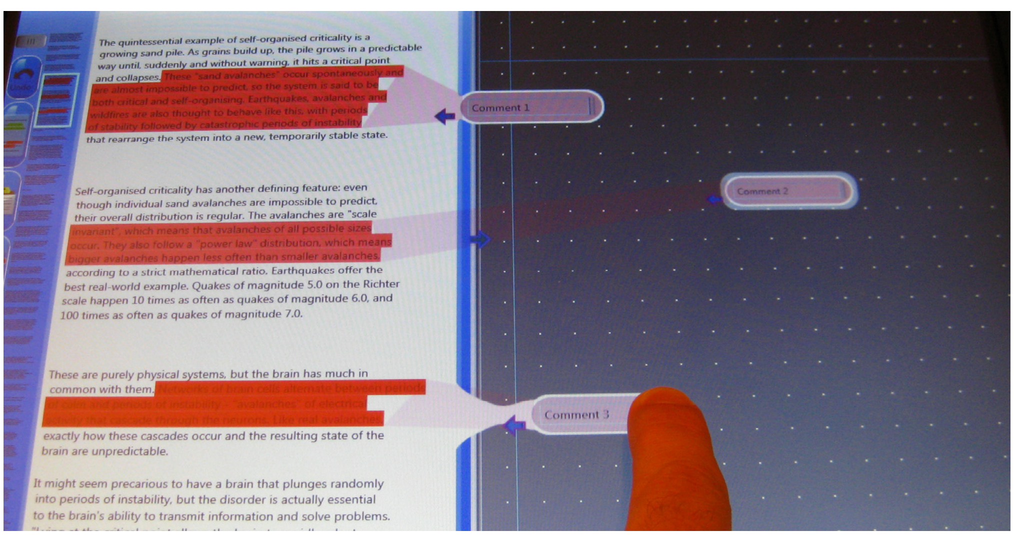


Figure 5.11 From top to bottom: a comment attached to the document, a comment pulled off the document and put in the workspace, and a comment being pulled away from the document.

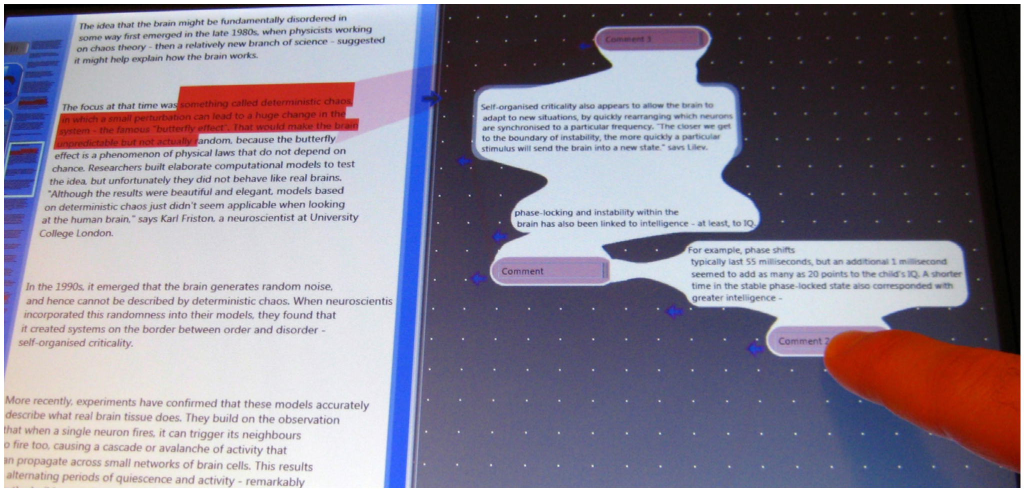


Figure 5.12. Comments and excerpts connected to form a large group.

fit to the space of a small margin, and are typically only able to refer to a single page of text at a time. While software such as Microsoft Word or Adobe Reader avoid some of these difficulties by letting comments expand a bit larger than they could with paper, they still largely follow paper's paradigm; their annotations are thus still limited to single referents, and control of the size or scale of annotations is very limited, so available space is easily consumed.

Following from my general design goals, I sought to provide a more flexible alternative in LiquidText. Text comments therefore act as attachments to the document itself (Figure 5.11). Like excerpts, they can be pulled off, rearranged, grouped with other items (including excerpts, as in Figure 5.12), and maintain persistent links back to the content they refer to. This allows the user to separate comments from their sources, aggregate and organize those comments, but without ever losing the link to their referents. This helps support separation, aggregation and retrieval of annotations, as suggested by [O'Hara and Sellen 1997; Renear, DeRose et al. 1999].

LiquidText also breaks away from a one-to-one mapping between content and annotations. Rather, a given comment can refer to any number of pieces of content across one or more documents (Figure 5.13). And since they maintain two-way links to their

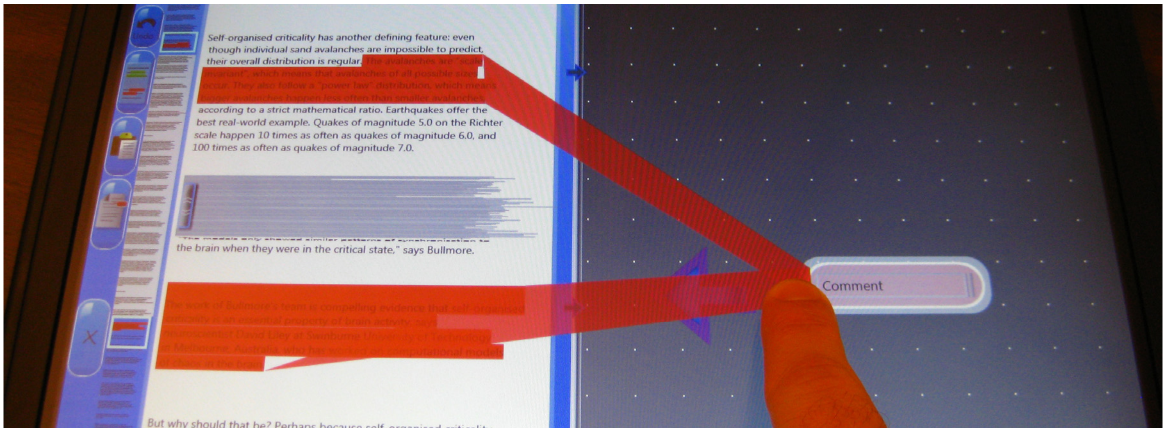


Figure 5.13. Comment's arrow-button collapses document to show both pieces of associated text.

referents, annotations can thereby act as persistent navigational affordances, letting readers freely create trails of links within and between documents.

The user has several options for actually creating a comment. First, they can select blocks of text in one or more documents or excerpts, and simply start typing. The system will create a comment object and, if all text selections are in the same object, the comment will be attached to that object. Otherwise, the comment will be positioned between the relevant objects. Alternately, the user can create an unlinked comment (i.e., a comment that does not link to any referent) by holding down one finger in the workspace while pressing the comment soft-button (Figure 5.1 A, 4th button from the top). This creates a comment in the workspace where the first finger was placed, and allows comments to act as very simple text editors, letting the user begin to work on the content creation tasks which often come later in the AR workflow. All comments then allow the user to type arbitrary amounts of text, and will increase in height as necessary to fit it. Finally, comments include a small handle on their right-end to allow the user to change their width (and thus the number of characters that can fit in each line).

Evolution

The commenting functions were early additions to the LiquidText system, and as such were available to evaluate in the formative study. There, participants did not identify significant problems with the functionality, but did provide feedback on how they used it. First, they revealed it to be used frequently, employed by 10 of the 18 users. The users also found comments beneficial for various purposes, particularly summarization and commenting, such as identifying contradictions and noting relationships to other projects. For the particular task we used in our formative study, users did not appear to have a need to associate comments with more than one piece of content; nonetheless, one user spoke more generally of this being one of the most valuable features of the prototype.

But since comments performed well in the formative study, and since they shared largely the same gestures as excerpts, I did not include them in the gesture refinement study.

Summative Study Feedback

Like excerpts, LiquidText's comments were important and frequently used, with summative study participants creating an average of 1.3 comments within the article being read, and 3.5 comments in the workspace. In LiquidText, comments served several purposes, but most frequently were used as headings to aid in the organization of excerpt groups in the workspace. Less frequently, they were also used to provide a summary or clarification of another piece of text, typically a part of the article being read. Least frequently, comments were used to record critical reflection and reactions to another piece of text—again, typically the article.

Subjectively, participants reacted positively to comments as well. Three users cited LiquidText's support for connecting comments together to form groups, and two noted that comments were easy to browse and search. Two participants also cited the freedom of comment positioning, in that comments can be removed from the document and arranged arbitrarily. Still, other aspects of comments were not found to be as useful. Virtually no users, for example, created comments referring to multiple sources at once. And very few users created comments referring to anything other than the original

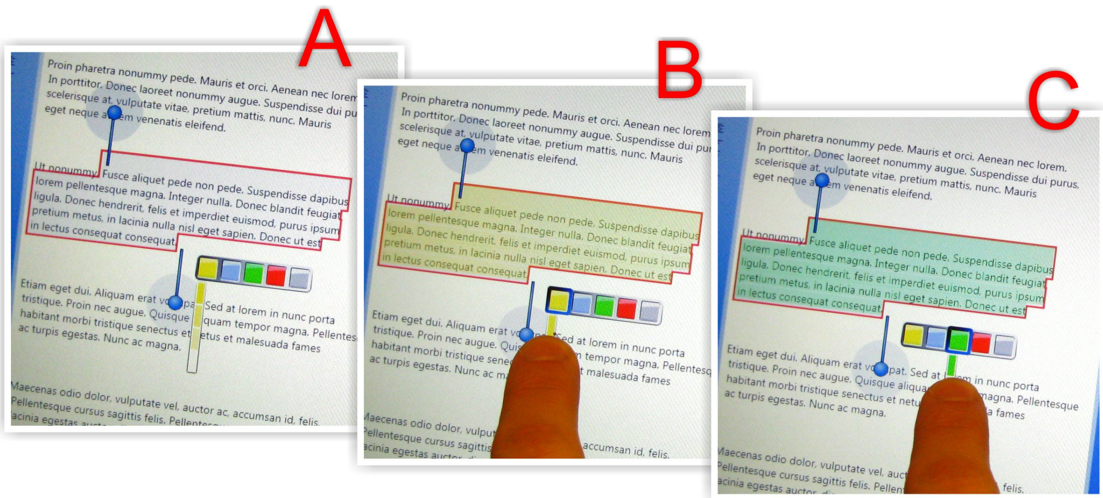


Figure 5.14. User highlighting text by A) selecting text, B) placing finger on desired color, and C) changing to other colors by sliding her finger horizontally.

document (such as an excerpt). Generally though, we see that comments were used frequently and received positively, and with virtually no negative feedback—even if, like the workspace, some of their more advanced features went unused.

5.4.2 Highlighting

Highlighting is one of the more common forms of annotation, raising the saliency of select passages of text by taking advantage of the pre-attentiveness of color. However, while highlighting is supported in many software systems (including Microsoft Word and Adobe Acrobat), a challenge of highlighting is that the desire to embellish many passages can easily lead to highlights losing their meaning, as I described in Chapter 3. Thus, I sought with LiquidText to give the user more control, making properties like color and intensity more readily adjustable.

To highlight text, the user starts by selecting a desired passage; this brings up a small, in-line color palette next to the selection (Figure 5.14 A). This palette provides two dimensions of control: color and intensity. The colors are shown horizontally, and the intensity is selected from a vertical bar shown underneath one of the colors. To select a highlight, the user begins by placing her finger on a color. This has two effects: 1) the intensity bar immediately moves beneath that color, and 2) the selected text is shown highlighted in the selected color. If the user is happy with the full intensity color, she can simply lift her finger and the interaction is complete. But if she would like a lighter variant, she can move her finger down the intensity bar to select one of three increasingly light versions of the color. And regardless of the intensity her finger is over, she can also move her finger left or right to keep the intensity constant but try out different colors (this is shown in Figure 5.14, where the user puts his finger down over light yellow, and shifts left to light green). Note that this interaction effectively lets the user select a color and intensity in either a single touch or, at most, a single movement of her finger.

To change a highlight, the user can double-tap a highlighted passage and the palette will reappear. To remove a highlight entirely, the user may select the lightest option on the intensity bar, which equates to "zero-intensity" and thus no highlight.

Evolution

Although highlighting is a common feature in document processing systems, in designing LiquidText before the formative study, my assumption was that with LiquidText's extensive functions for content extraction, commenting, and layout, highlighting would be of little value. As such, I included only very simplistic, single-color highlighting in the earlier prototype. The user could thus highlight text (in yellow) by selecting a passage and pressing a highlight soft-button, or by holding down the highlight soft-button while dragging her finger over the desired passage.

The formative study showed my assumption to be partly true—while 9 of the 18 participants used highlighting, those users who tended to make more use of LiquidText's excerpt and commenting functions also seemed to make less use of the highlighting. One participant even described highlighting as possibly redundant with the system's other functions—especially excerpts.

Nonetheless, some users' do have a strong preference for highlighting, prompting me to design an alternative. In this version, the user had a fixed color palette in the corner of the display where she could select a desired color. To highlight text, she would select a desired passage, put her finger on top of it, and drag up; the distance she dragged up controlled how intense the color would be.

While this did provide control of both color and intensity, it had problems that emerged in the gesture refinement study. Particularly, users had difficulty performing the drag-up gesture over text passages; they tended to make a small, fast stroke rather than a controlled, long stroke. They also had difficulty understanding where they were supposed to start and complete the stroke. As such, my design team and I redesigned the interaction to its final form, providing a more efficient as well as more intuitive mechanism for highlighting. As shown in Figure 5.5, users were indeed more comfortable with this later interaction.

Summative Study Feedback

Highlights were among the most frequently used functions in LiquidText, and were employed by 9 of the twelve participants, with an average of 8.4 highlights per user (among those who used highlights at all). Likewise, LiquidText's focus on providing easy control of color and shade appears to have been taken advantage of, since participants used an average of 1.4 different highlight colors/shades in the control, versus 2.8 with LiquidText.

In the interviews, participants reported using highlights as a way to raise the saliency of text, which was in line with my expectations. This included highlighting material because it seemed important to the author or because it fit with the task description. Some participants additionally employed highlights in tandem with excerpts. For example, one described using highlights to identify material the author seemed to find important, while using excerpts to identify material that represented the key ideas of the article. Another participant first highlighted the article, and then created excerpts out of the most important highlights.

The interviews also revealed subjective reactions to the highlighting functionality, and found participants expressed little excitement about the functions, but also few complaints. Two participants did, however, question the breadth of colors and shades, with one arguing that he would not want different shades of the same color since he might have trouble selecting the same shade consistently. Another simply felt one color was enough. However, these reactions appear to represent the minority since users did on average tend to employ more highlight colors/shades with LiquidText than the control. So generally, while its breadth may have bordered on excessive, highlighting largely appears to have fulfilled its purpose of providing a simple, flexible way to raise the saliency of text.

5.4.3 Displaying All Highlights

Not surprisingly, one of the key requirements for annotation is that it be easy to navigate through and retrieve [Renear, DeRose et al. 1999]. As dicussed in chapter 3, my own research confirmed this, with the difficulties participants expressed with retrieving comments and highlights. As such, LiquidText includes features to help users re-find highlights. The first of these is to show which areas of the document are highlighted in the preview pane. The second is to actually display all highlights together.

In designing this latter feature, I considered several approaches with different tradeoffs. One approach was a simple list (akin to [Marshall, Price et al. 2001]), where LiquidText would provide a linear list of all the highlights in a scrollable form. This approach is conceptually simple, but can easily take the highlights out of their context, forcing the user to check their contexts one-by-one when needed. Additionally, such a function would require a large amount of display space to work well, and display space, even with the 17” monitor, was at a premium.

An alternate approach was to distort the document itself, such that only the highlighted areas were visible. The advantage of this approach is that it can be more

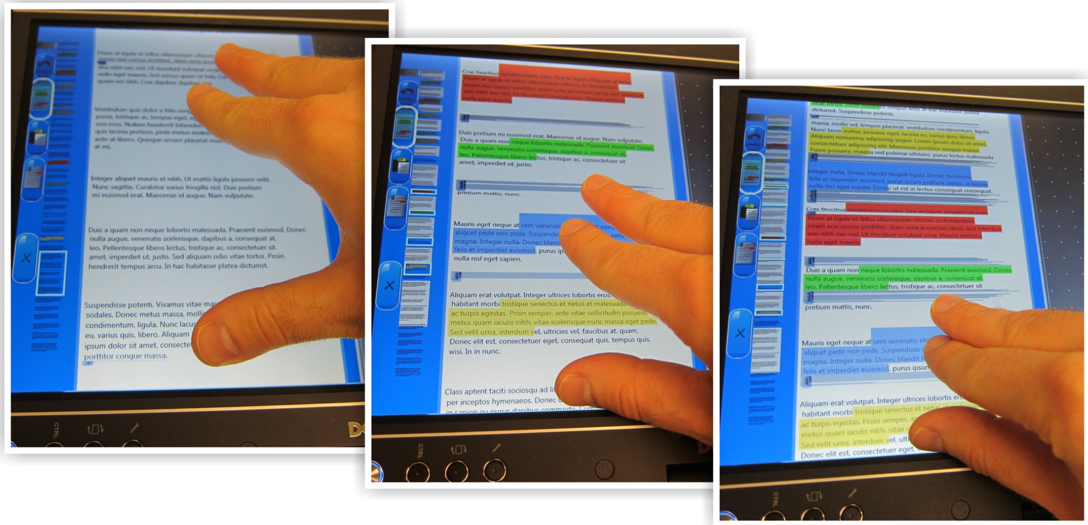


Figure 5.15. Three successive stages of collapsing a document to see all highlights. Note that the intermediate stage allows more context surrounding the highlights to be seen.

flexible, for example, showing additional document context inline, and using the document itself to display the highlights rather than requiring an additional window. The main disadvantage, however, was that the user has no separate object containing the aggregation, making it ephemeral. Ultimately, I nonetheless chose this approach because it fits with the general idea of LiquidText of affording the user more control over the arrangement of content in their document. But to mitigate the problem of ephemerality, it was essential that this aggregation could be created quickly and easily.

The resulting interaction is based on the collapse idea described above. To see all highlights, the user first taps a highlight aggregation soft button (Figure 5.1 A, 2nd button), which enables highlight aggregation mode and displays a simple animation explaining what to do next. After this, the user performs the same 3-finger collapse gesture described above. LiquidText then collapses the document, but with two differences: 1) only unhighlighted text is collapsed, and 2) text is collapsed much faster, so that a full sweep of the hand will collapse the entire document (except, again, for the highlights). The result is that the user can see only the collapsed text (Figure 5.15).

As a result of using the collapse concept, this interaction gives the user precise, flexible control of the aggregation process, specifically: 1) The further the user pinches, the more of the space between the highlights is removed, letting the user select how much context is seen around highlights. 2) The direction of the pinch (top finger brought to bottom or vice versa) controls whether the part of the document above or below the user's fingers is collapsed. This was necessary for scenarios where the user might wish to aggregate highlights in part of the document, while reading another part normally. And 3), this operation works internally by performing a "normal" collapse operation (described above) between each region of highlighted text. Thus, a user can view additional context around a single highlight just by manually expanding the collapsed areas around it.

When the user is finished viewing their aggregated highlights, she may uncollapse the content using the reverse of the collapse gesture, or she can just tap the preview panel to make the entire document fully uncollapse. She can then tap the highlight aggregation button again to exit aggregation mode.

Evolution

As noted above, the highlight aggregation functionality was a result of the formative study, and so not evaluated therein. The gesture refinement study, though, prompted considerable changes to the user interface for this function.

In its initial form, the highlight aggregation was designed as just one of several functions to provide *in*direct manipulation of objects based on color—i.e., having all objects of a given color behave in a certain way. As such, LiquidText at that point had a fixed color palette; each color could then be "opened" to access various controls, one of which was filtering. One could enable this filtering for one of more colors, and then perform a collapse gesture; this would then have basically the same effect as in the current system, except it would only ensure visibility of the selected colors. Users however, reacted very negatively, seeing this process as requiring far too many steps. Moreover, they found it highly counterintuitive that enabling the filtering for a given color was not enough to make it auto-collapse to show just the highlights of that color. That is, the fact that they *then* had to perform the collapse gesture was seen as confusing.

In response, my design team and I explored numerous alternatives, including different gestures to perform the collapse, ways to avoid the modal-ness of the interaction, ways to offer additional feedback, etc. Out of this theoretical and empirical trial and error came several points. The first was that the collapse gesture itself was viable for this interaction; even simpler alternatives that left the user with less control did not seem to improve the overall user reaction. Second, while the modal-ness of the interaction is generally not advisable, we needed *more* complex gestures to make it non-modal, and those were far more problematic. Third, while the modal interaction was tolerable to users, more feedback was required to remind them of the steps of the interaction, which took the form of the explanatory animation noted above.

The final user interface for highlight aggregation as described above never earned high praise from users, as seen in Figure 5.5, but did ultimately receive a better response with later iterations. Also, qualitatively, users in the gesture refinement study generally appeared to struggle less with the final design, and appeared to be more readily able recall and perform the interaction.

Summative Study Feedback

Much like the collapse functions, the highlight aggregation was used by only two participants, and thus elicited little overall feedback. Still, those two participants did speak positively; with one calling it “amazing” and another saying it is an advantage over paper. But by contrast, a third participant reported trying to use the highlight aggregation but failing. He tended to highlight many pieces of text in close proximity, and so collapsing only entire *lines* of text did not elide enough content to significantly aid in seeing the highlights in aggregate. However no participants reported having difficulty

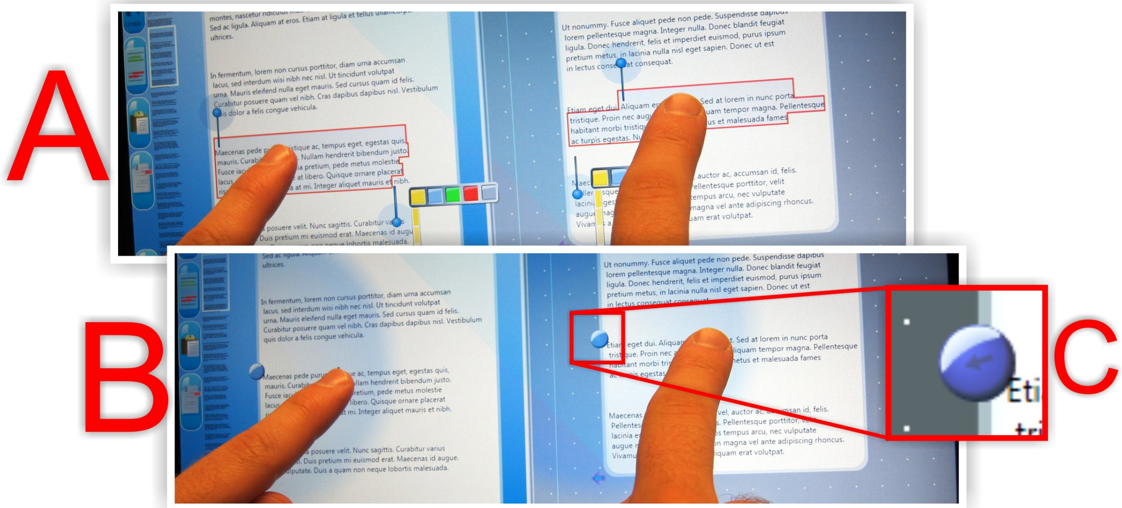


Figure 5.16. A) The user is about to touch two selections to create a link between them. B) Immediately after touching them, a ripple animation is displayed and link buttons are created next to each selection. C) A close-up of one link button.

with the aggregation gesture itself, suggesting that the iterative design process did result in an improved aggregation interface.

5.4.4 Direct Linking

As discussed above, LiquidText allows comments to link to multiple objects or pieces of text at once—effectively enabling comments to act as links across one's corpus of materials. And while participant reactions to this functionality were positive, there was an interest in linking text *without* needing to involve a comment—i.e., the added semantic expressiveness of a comment was superfluous. This interest came partly out of the feedback I received from participants after using the prototype, and also partly out of the design workshops, where some participants included direct linking functionality in their mock-up designs.

As a result, I added a direct linking function to LiquidText to allow the user to easily create an association between two arbitrary passages of text. To create a link, the user begins by selecting the two desired pieces of text (which can be in one or more documents or excerpts). Taking advantage of the parallelism of multitouch, the user then just touches both text selections at once, and a link is created between them (Figure 5.16

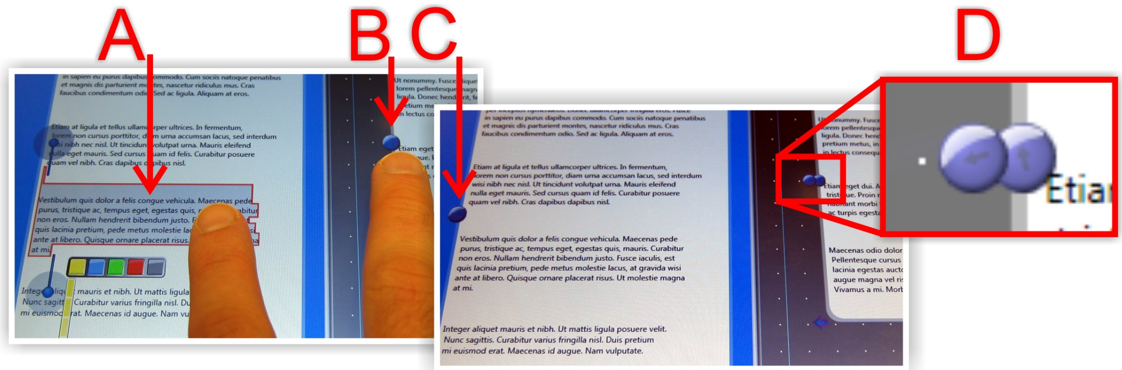


Figure 5.17. User linking an existing link button (B) to an additional text selection (A). As a result, the link button is displayed as two small overlapping link buttons (D). The new selection also gets its own link button (C) which points to (D).

A). In creating the link, the system immediately adds link buttons next to the text selections and displays a ripple animation emanating from each of the two buttons in order to draw the user's attention to their addition (Figure 5.16 B).

Analogous to the arrow buttons used elsewhere in the UI, touching a link button causes documents and other relevant objects to scroll or move such that the content associated with the opposite link button is displayed. To give the user a small cue of what a link button points to, each button has a faint arrow within it that points to its referent (Figure 5.16 C).

Beyond connecting only two pieces of text, LiquidText allows the user to add multiple referents to a link button—i.e., a link button can point to several pieces of text. To do this, the user chooses an existing link button and selects an additional passage of text she would like that link button to point to. She then touches the selection and the link button simultaneously, and another link is created (Figure 5.17). Now, the selected link button points to two pieces of text. To indicate this to the user, the button becomes a *composite* link button, and is shown as two small buttons partially on top of one another (Figure 5.17 D).

Touching a composite link button will cause the relevant documents or other objects to scroll, move or collapse as needed to make all referents visible (Figure 5.18, A). Alternately, the user can access each referent individually by "expanding" the

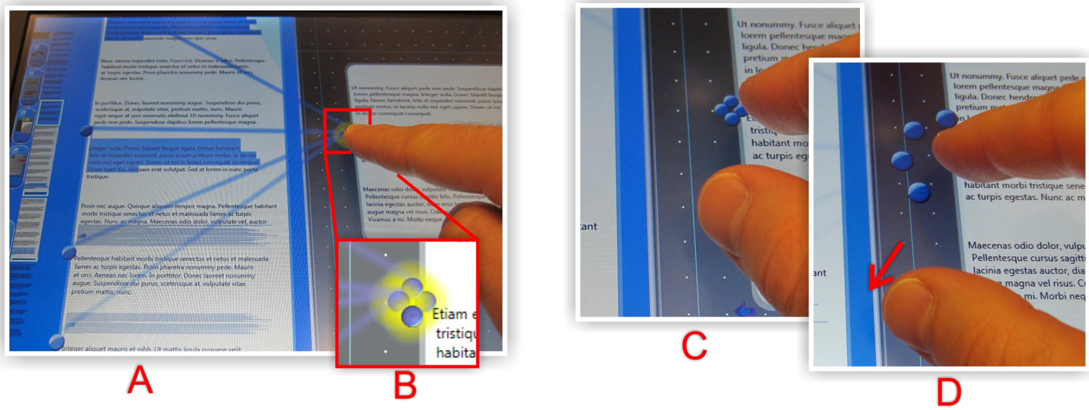


Figure 5.18. A) Touching a link button that points to 4 referents (close up of button (B)). User can expand the link button (C, D) so she can access each of the referents individually.

composite link button. This is done with a simple two finger stretch gesture performed roughly over the button (Figure 5.18 C, D). As a result, the smaller buttons displayed within the composite expand and separate so that the user may access each one individually. As one would expect, the composite link button can be un-expanded again by reversing the pinch gesture used to expand it.

While I expected that the linking functionality would form a valuable component of LiquidText, I did not include it in the gesture refinement study. My decision was based on the limited time I had with study participants, and my conclusion that the linking was a *valuable* but not essential feature. Thus I concluded that the design was adequate as-is for the summative study.

Summative Study Results

Like several of LiquidText's other more advanced functions, linking was not frequently used in the summative study, but was still employed by two participants. Both used the linking in similar ways, connecting two disparate areas of the document they were reading. The interviews, however, did not reveal significant additional details, except that one participant found the linking interaction difficult to remember.

The interviews did, however, point toward a different implementation of linking. As it stands, linking operates on the level of text—a string of text on one object points to some other string of text. But two participants described a more flexible, loose approach to linking, where the links could operate on higher or lower levels. They wanted to be able to indicate, for example, that two entire excerpts related to one another, or that several groups of excerpts were related. In essence, it appears that they sought a feature analogous to drawing arrows from one area or object of the workspace to another. While it is unclear whether the absence of this functionality inhibited more people from using the linking function, it does point to an interesting direction for expressing relationships in future revisions of the design.

5.5 Word Processor

Fundamentally, my research with LiquidText focused sharply on active reading, and not on the broader document workflow that may accompany reading. But in order to evaluate LiquidText in the summative study described in the following chapters, it was necessary to enable a larger document workflow. In this case though, the workflow is extremely simple, and involves just reading a document and then writing a response. LiquidText itself was thus largely adequate, but still required functionality for composing a written response.

The most natural way to provide this support is simply to offer the user an off-the-shelf word processor—as I did in the formative study. Recall that the formative study, too, involved the participants reading a document and writing a response. As such, participants had a hybrid computer setup, including both a touchscreen unit running LiquidText, and a mouse/keyboard unit running the word processor. This resulted in problems however, as participants had to repeatedly switch between a mouse and a touchscreen. Participants did not provide details on their difficulties, but part of the problem was likely the ergonomic differences: a touchscreen should be angled and fairly close to the user, whereas a monitor used with a mouse should usually be vertical and further back. And differences in viewing properties, like angle, was a difficulty participants cited with paper. Likewise, LiquidText is asking the user to work in very

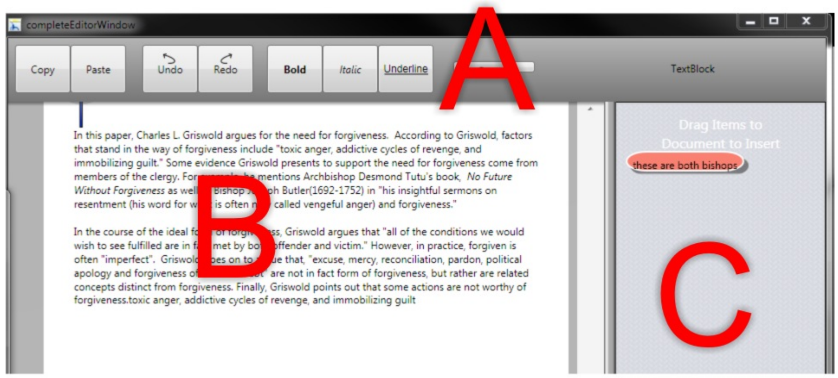


Figure 5.2. ***(Repeated here for convenience)*** **Overview of the word processor linked to LiquidText: A) Button bar, B) document being written by user, C) area where text selected in LiquidText appears.**

unfamiliar ways and to develop new habits, but providing a mouse raises the expectation of working in mouse-oriented ways, which are both infeasible and suboptimal in LiquidText.

As such, I developed a very simple, entirely touch-based word processor to be used with LiquidText. The word processor itself contains three pieces, as shown in Figure 5.2: A) a toolbar containing basic formatting options, cut/paste and undo; B) the document itself where the user can type as in any text editor; and C) a text insertion pane which allows text to be copied from LiquidText to the text editor. I discuss each of these in turn.

As noted above, this word processor is only intended to support writing text responses, rather than page layout or other advanced tasks. As such, the user is only given three basic formatting options: bold, italic, and underline. The toolbar provides large buttons to toggle these options, but standard Windows keyboard shortcuts work as well (Ctrl+B, Ctrl+I, and Ctrl+U, respectively). Additionally, the toolbar contains buttons for unlimited undo/redo, as well as copy and paste to/from the standard Windows clipboard.

With respect to the keyboard, the document itself behaves largely as in other text editors. However, the document also supports touch—particularly, a small subset of the same gestures used in LiquidText, including: 1) the user can use two fingers to pinch/stretch the document to zoom in or out; 2) the user can drag a finger over the

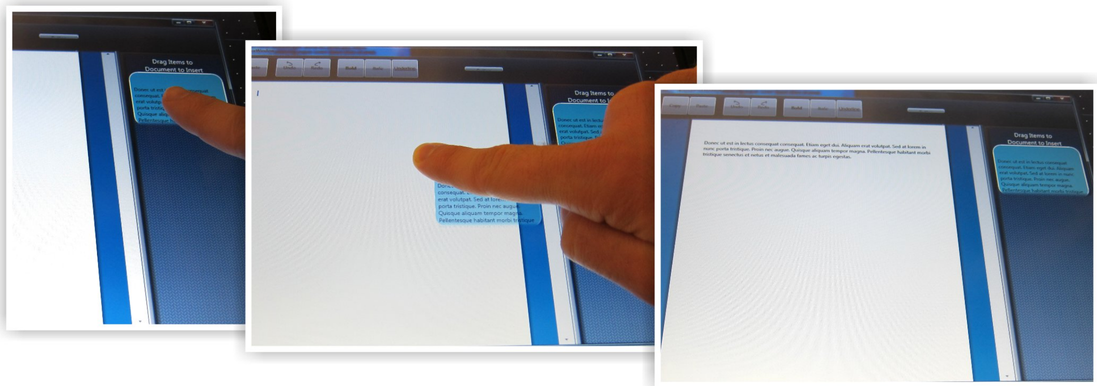


Figure 5.19. Three steps of the user dragging text from the insertion pane of the word processor into the document which the user is writing.

document to scroll it; 3) the user can select text through the same tap and drag gesture, and with the same adjustment handles, described above; and 4) the user can move selected text to a different part of the document by dragging it with one finger.

The third component of the word processor is the text insertion pane. When the user selects text or touches a comment or excerpt in LiquidText, a copy of that text is placed in the word processor's insertion pane. From there, the user may drag the text into the document they are writing. This enables the user to copy text quickly and efficiently from the LiquidText environment into their response document (Figure 5.19). To minimize visual complexity and confusion, the insertion pane only shows the current text selections, if any, and the one most recently touched comment/excerpt.

5.6 Functionality not Included

In designing LiquidText, there were a variety of areas of functionality that I explored, but which were not included in the final system. The development of these functions ranged from simple design sketches to prototype implementations. But from the perspective of providing an adequately rich environment to evaluate my research questions, none of these functions were strictly necessary, though the particular reasons they were rejected ranged from technical infeasibility to the time constraints on the implementation phase of the project. In this section, I will discuss two of these functions to illustrate some of the directions LiquidText could be taken in future extensions of this research.

5.6.1 Dog-earing.

Bookmarks meant for short-term, transient use are an important part of active reading navigation [Askwall 1985]—as when holding a thumb at a particular page while browsing a book. Such "ephemeral" ways of book-marking are often used when rapidly flipping between regions of a text. In addition to offering several options for persistent book-marking (via excerpts and comments), LiquidText includes a way to create a form of "transient" bookmark. As explained above, this is done by simply holding one's finger on a piece of text—this will keep the selected text visible while one performs other navigation operations (as LiquidText will collapse the document so all selected areas remain visible). But to give LiquidText a richer bookmarking system I explored a second way as well, with behavior more like that of a traditional bookmark, which I call "dog-earing." This feature was focused on scenarios where the user would create and then refer to a bookmark within a short period of time, and is intended to support near effortless creation and deletion of the bookmarks.

This dog-earing interaction was modeled on its paper counterpart: users associate a finger with a given document location, much like putting a thumb on a book page allows a user to rapidly flip between places in the text. In a prototype implementation which I developed, creating one of these bookmarks entailed the user putting down a finger in a special "dog ear region," displayed in the lower-left corner of the monitor. This creates a small orb under her finger corresponding to the current state of the document (including where it is positioned, whether it is collapsed, etc.). The user can then navigate as she wishes using the other hand, while keeping her finger on the orb to "save her place." To return to the captured state, the user simply lifts the finger off the orb and LiquidText returns to that state while the orb fades out. Taking advantage of multitouch input, the user may do this with more than one finger at a time to capture and flip between several layout states. To discard an orb without using it, the user drags it away from the dog-ear region before releasing it.

The choice of gesture here enforces transience, in effect: the bookmark will be kept for only as long as the user's finger is in place (just as with a page in a book). Bookmarks held in this way vanish as soon as the finger is lifted, meaning that the user is freed from having to explicitly delete the bookmark. (Other forms of book-marking can be used when a user wishes to create a more persistent placeholder in a text.)

Status

Dog-earing is one of a handful of functions I developed for LiquidText which were disabled for the studies. In this case, the reason was that the touch-sensing hardware was not adequate to support the dog-earing gesture. With the touch tablets that I used in the formative and gesture refinement studies, the user could not hold a finger in the same place for an extended period of time, or the system would "get used to it" and rapidly alternate between detecting and not detecting the finger. As a result, it was not feasible to iteratively refine this interaction on the same schedule as the rest of the system. Since I believed this interaction would have required significant additional refinement, and since there were already other functions that achieved a similar purpose, I did not enable the dog-earing in the version of LiquidText used in the studies.

5.6.2 Pen and Touch Support

While LiquidText's existing annotation system offer extensive control over the visual properties of comments and highlights, and provides other annotation-like functions such as linking, the formative study participants expressed an interest in even more idiosyncratic annotation. In particular, participants indicated that they would like to be able to annotate using their own handwriting and semiotics using some type of pen interface. This was likewise demonstrated in the design workshops, where many of the participant mockups included pen or pen + touch as an input modality.

This interest in mixed-modality input was part of a general recognition that neither a pen nor a finger is really ideal for all input. Rather, manipulation operations, such as moving, resizing, etc., were well suited to fingers, and drawing ink strokes and small-scale manipulation tasks, such as selecting text, were seen as better suited to a precision instrument like a pen. As such, I explored a simple pen-based inking function for LiquidText and provided a prototype implementation.

My broader vision for ink support in LiquidText was to allow ink itself to be flexible, and to acquire simple semantic properties, in some ways akin to [Elrod, Bruce et al. 1992]. Particularly, the user would be able to draw within and *across* documents, using ink to express intra- as well as inter-document relationships. The ink would be intelligent, and stretch/distort as necessary when documents were moved, resized, collapsed, etc. Selected ink strokes could then be grouped by the user into an object, which would possess the same semantics as a comment, and thus be able to link to referents, attach to other comments, etc.

In order to explore and experiment with this concept, I brought another student into the LiquidText project to implement some basic aspects of this vision using the pen + touch support on our Dell XT2 touchscreen tablets. The result was support for simple inking within objects. The user could draw ink freely within any document or excerpt and the software would deal intelligently with ink drawn outside any object (by trying to tie it to the most likely intended object). Notably, ink drawn on a document could distort appropriately if the document was collapsed.

Status

After completing this prototype implementation, I ultimately decided not to include ink support in LiquidText for two reasons. First, the technical challenges of

Function	Active Reading Process	Changes after formative studies
Scrolling/dragging	Navigation	Unchanged
Preview pane	Navigation, especially providing visual metadata	Unchanged
Rescale objects	Layout	Simplified gesture, merged with rotate gesture
Rotate objects	Layout	Simplified gesture, merged with rescale gesture
Manual collapsing	Layout and Navigation, especially parallel viewing	Modified gesture after formative study
Indirect collapsing	Layout and Navigation, especially parallel viewing	Minimal changes.
Fisheye workspace	Layout and Navigation, especially maintaining awareness	Resulted from formative study, refined in gesture study
Text selection	Enables excerpts, annotation	Refined to simpler, more reliable gesture
Excerpts	Extracting content	Simplified creation gesture
Comments	Annotation	Minimal changes
Links between excerpts/comments and their contexts	Navigation, especially traceability of excerpts and retrieval of annotations	Included in initial build, unchanged thereafter
Connecting excerpts/comments	Layout, especially for organization and idiosyncrasy	No change
Resizing/rotating objects	Layout, annotation, especially for idiosyncrasy	Combined and simplified gestures
Highlighting	Annotation	Added colors and in-line highlight control
Showing all highlights	Annotation retrieval	Resulted from formative study, simplified and added more feedback in gesture study
Direct linking	Navigation	Resulted from formative study
Word processor	Required for study	Resulted from difficulties in formative study setup

Navigation Layout Extracting content Annotation

Figure 5.20. The core functions of LiquidText, color coded broadly by active reading process category.

handling large, complex ink diagrams efficiently (especially during complex transforms like collapsing) would have required more time and resources than were available. Second, I did not have a viable pen + touch platform. The Dell XT2 tablets could detect both, but not simultaneously and not reliably. Second, the larger 17” touchscreens I intended to use in the summative study could not support pen input at all. Thus, pen support was not included in the final version of the system.

5.7 Implementation

In this section, I will briefly elucidate LiquidText’s architecture and point out several potentially noteworthy aspects of the development. I will also briefly discuss the hardware platforms on which LiquidText runs.

5.7.1 Architecture

The LiquidText application consists of about twelve thousand lines of C# and XAML code. It runs in the .Net framework, and relies on the Windows Presentation Foundation hardware accelerated rendering for all of its graphics. I developed it using Microsoft Visual Studio and Microsoft Expression Blend.

As the Net framework does not offer substantial support for multitouch, the core of LiquidText’s architecture is largely a multi-point input manager. At its heart, this includes abstract classes that know how to respond to parallel touch input messages, as well as various touch data structures. On top of that, the core infrastructure includes functions for identifying the targets of touch input and passing the touch messages appropriately.

An interesting challenge that comes with processing multitouch input is that relevant touch data may not be physically directed at the object it pertains to. For a simple example, in one of the earlier iterations of LiquidText, moving a finger over a block of selected text would normally drag the whole document; but if a thumb is down somewhere else on the touch-screen, it would be an indication to duplicate the text instead. For this kind of situation as well as others, it is important for multitouch-enabled objects to be able to issue queries as to the state of other potential inputs. For multitouch systems that support arbitrary numbers of inputs, like LiquidText, procedurally coding such queries is impractical. Thus, LiquidText includes a simple querying scheme, where objects can pass universal and existential lambda-calculus predicates that are used as the basis for analyzing the current set of touches. For example, when an object is informed that it has been touched, it can issue a query asking whether there exists a forefinger sized object touching it, and whether there does not exist a thumb-sized finger that is touching anything else. If there are touches that meet all positive criteria and no negative criteria, they are returned to the object for further processing. While somewhat arcane, this system is flexible enough for handling a very wide range of multitouch gestures.

On top of the touch management infrastructure are various concrete, multitouch-enabled visual objects. The most complex of these are used to represent document objects, such as original texts, excerpts, or comments. The base class for all of these contains custom text layout code needed to manage visual effects such as text collapsing.

5.7.2 Fisheye Workspace

Among the more complex implementation tasks in LiquidText was the Fisheye Workspace. The challenges of this component emerged from two causes. The first is the complexity of generalizing the basic LiquidText interactions so they behave correctly and predictably with a non-linear, non-uniform, and effectively arbitrary mapping between the coordinate system in which the objects reside and the user's display coordinate system. Complicating this task immensely was that I could not simply use a pixel-to-pixel transform to do the mapping, which would have been possible by simply convolving the output bitmap with a transform matrix, and using the inverse transform matrix for input touch points. Rather, this fisheye effect had to operate on the object level instead of the pixel level, and so each object needed to be transformed individually. Thus, the many factors involved in an object's on-screen position and size (including the position/size of its chain of parent objects, any manual resizes/moves the user has performed, and any transforms coming from the fisheye distortion) had to be kept carefully disentangled. For example, recall that the fisheye distortion will increase the sizes of some objects while reducing the sizes of others. Now let us say the fisheye distortion caused parent object O to increase in scale by 10%, that would automatically propagate to O's children, increasing their size by 10% as well. But suppose that the fisheye transform was such that O's children are supposed to actually *shrink* by 10%. As a result, the fisheye transform would have to actually shrink O's children by *18%*. Therefore, the fisheye algorithms cannot blindly apply their transform, but must compensate for, and work with, the other relevant transforms involved in an object's rendering to achieve the final desired effect.

The other, lesser challenge for the Fisheye Workspace was identifying the proper fisheye transform. For this, I first explored a simple, piecewise linear transform (in radial coordinates). Its abrupt changes with increasing *r* were somewhat jarring, though, so I changed to a non-linear function. Thus, the final fisheye transform is a piecewise, continuous (equal at the first derivative), 2nd degree polynomial equation.

5.7.3 Hardware Platform

As I allude to above, modern multitouch sensors generally involve significant tradeoffs. Highly accurate multitouch generally involves very physically large devices (e.g., [Han 2005]) based on rear projection, and would be challenging to use in a desktop application. Small systems like the iPad can be quite accurate as well, but their projected capacitive sensors do not scale well and thus this high accuracy is generally not feasible above their 10" or so scale. Thus, the best compromise I could find to have a large display, and a manageably-sized overall device, was the 17" projected capacitive touch sensor from N-Trig[5]. But as I noted, projected capacitive touch sensing becomes more problematic as scale increases. The 17” sensors are thus highly sensitive to environmental electrical noise, such that if certain fluorescent lamps are in their vicinity they will not reliably detect touches. Likewise, this class of sensors appears to require a very good ground, and can become extremely inaccurate if run off of batteries instead of being plugged into an electrical outlet.

Most of these problems, though, could be managed by controlling the experimental environment. The more serious problem with the 17” projected capacitive sensor was that, at least in our configuration (described in the following chapter), they would simply become less accurate with running time. The reason for this was never established, but leaving the computer running for several hours would result in almost unusable touch sensing. Worse yet, they would sometimes have low accuracy for a given user for no clear reason at all. Still, these problems were typically rare and not severe enough to threaten the study itself and so I used these 17” sensors for the summative study.

5.8 Discussion of Design

Beyond my research questions per se, a core contribution of this research is in the design of LiquidText—both in specific areas of functionality as well as in the underlying approach to representing content. In this section, I consider some of the implications and possible lessons for the broader design community emerging from LiquidText. I begin by exploring aspects of the system’s feature set and design approach that could be applied in other domains, and conclude with several lessons learned through the design process.

[5] www.n-trig.com

5.8.1 Generalizable Ideas

Among the core ideas behind LiquidText's design was to offer the reader flexible, rich ways to annotate. That is, to be able to express a wide variety of thoughts and observations into the reading medium. A core part of this was in LiquidText's approach to comments. While commenting is used broadly across document preparation and reading software, it faces various constraints; notable among these is that it is extremely unusual for comments to support multiple referents at once. LiquidText, by contrast, allows comments to refer to an arbitrary collection of text passages, even spanning multiple documents. In this way, comments can act as links, joining different documents or document areas together. Such an approach to comments could be of value in various document annotation situations where users identify non-localized relationships, such as observing several related web pages or spreadsheet cells. Still, in the summative study, comments were not often used to connect disparate content, so additional study may be required to fully understand the domains and tasks where they offer the most benefit.

But perhaps a more significant characteristic of comments in LiquidText is that they can be manipulated independently of the document. This is a significantly different model of comments from that used in most annotation software, which either places comments at fixed locations, or largely constrains them to the vicinity of the document. Yet such flexibility appeared to be of value to participants in the summative evaluation, 58% of whom did pull comments off of their documents and into the workspace. As participants explained, part of the value of this was in simply being able to see all comments at once—by itself, a common feature in many document readers. But participants also noted that they liked the ability to connect comments to other objects, and to organize them spatially. Likewise, a similarly flexible model of comments might be of value in other domains that require annotation, especially those requiring critical thinking, analysis and other tasks found in active reading.

Beyond the annotation of text, LiquidText's functions for the visualization and arrangement may likewise have potential to be applied more broadly. The two most iconic and well-received of these functions were excerpts and collapsing. While they behave very differently, both functions enable users to conceal text deemed irrelevant, as well as view disparate areas of a text in parallel. Excerpts arguably went further in this task, allowing the user to construct a completely different organization of a source document, including hierarchies and groupings, without losing their links to the source text. By contrast, collapse more strongly preserves the document's linearity, but does not provide as flexible a visualization. But both of these could be applied beyond the bounds of LiquidText and even beyond active reading. The idea of collapse could be employed in many visualization tasks where data populate a 2-dimensional plane, especially if one dimension is of fixed size, as with written text. Collapse is especially appropriate for transient views, where the user is performing only a brief sub-task. By contrast, LiquidText's excerpts are more appropriate as a longer term visualization, and tend to make sense in many domains, and without as many assumptions about the spatial structure of the source content.

And similarly to collapse, the magnifying glasses likewise potentially offer a powerful visualization technique for exploring two-dimensional data. Of course, lenses and fisheye views are not new in themselves, but the multitouch interaction used to control them in LiquidText raises the potential for more natural, intuitive ways to control this type of distortion. Still, the summative study results suggest that the magnifying glasses were not extensively used in LiquidText, and more research may be needed to fully understand how best to apply them.

General Approach

Underlying LiquidText's specific interaction techniques, however, is the notion of flexible, high degree-of-freedom representations of content, combined with comparably rich input devices to take advantage of them. With LiquidText, I sought to apply this in the context of active reading, but the approach may be advantageous in other domains as well. As discussed in Chapter 7, summative study participants tended to respond most favorably to the aspects of LiquidText that provided more freedom and control than they would have in more traditional media. This included control over what areas of a text were visible, the positioning and organization of content and annotations, flexibility in navigation, and the like. And while the summative study suggests the value of the ideas behind LiquidText, the formative study suggest its breadth of applicability. As discussed in Chapter 3, formative study participants showed that tasks like email, spreadsheet use, and manipulating presentations often intersect with AR, and share many of its challenges—like visualizing disparate pieces of content in parallel. As such, many tasks beyond traditional text document consumption may benefit from an application of the same principles behind LiquidText.

LiquidText's use of multitouch additionally turned out to be an essential component for realizing the design goals of the system. Interactions involving many degrees of freedom, such as the fisheye lenses and collapsing, depended heavily on this type of input—they otherwise would likely have required many separate controls or highly serialized interaction sequences. The value of this input model was also supported in both the summative and formative studies, with many users seeing touch as natural and very direct. Equally important, participants who discussed the issue mostly felt multitouch was appropriate for this domain of knowledge work, with the exception of certain interactions, such as text selection. Likewise, participants rated the ease of recalling and performing LiquidText's gestures at 4.6 and 4.2 out of 5, respectively, suggesting that even this atypical multitouch gesture vocabulary posed little difficulty. Generally then, multitouch was well received as an input modality for professional knowledge work, and thus raises the possibility of applications in other knowledge work domains.

5.8.2 Lessons Learned

The conception and iterative design of LiquidText offered a variety of lessons. In this section, I expand on two of these: the tension between flexibility and structure, and the importance of UI feel.

Design processes can often include tensions between competing goals—such as size and weight versus performance in a computer. And in LiquidText, the design goal of flexibility came into tension with usability. In a sense, the use of multitouch was a mechanism to resolve one instance of this, as multitouch offered the potential to let the user control many degrees of freedom without sacrificing a natural, memorable interface (independent of the issue of efficiency). But in other instances, such a resolution was less apparent, as with the balance of flexibility versus structure. That is, I observed a tension between the flexibility offered to users in performing their task, versus the amount of predefined structure provided to simplify the task. This was evident in several places, such as the mechanisms for grouping objects together. As described above, more structured groups can use space more regularly and efficiently, and be simpler to use. But such structure can also be presumptuous about the ways the user would want to organize their materials, and can deprive the user of the idiosyncratic visual cues and personalization that come with freer organization.

While my solution to this and similar tensions was to provide more flexibility, another option could have been to make structure optional. But even approaches to making structure optional sometimes seemed to impose structure. For example, my design team and I considered allowing users to name object groups, but to make the feature visible required a default title already in place; but any meaningless, default title (e.g., "Group 1") would be effectively begging the user to rename it, meaning that the choice of whether or not to name is no longer as optional. Given the goals of LiquidText, I generally opted for flexibility, but a possible opportunity for future research is to explore the question of how best to balance these competing ideals.

The second lesson concerning the design of LiquidText is in the value of reaching beyond mere productivity. While most participants discussed LiquidText in largely goal-oriented terms, a minority across the summative and formative studies specifically noted LiquidText's impact on affect. For example, a formative study participant described the system as having a bubbliness, and reported liking its "personality." A summative study participant said LiquidText is "just really cool and a lot of fun," and likened its use to playing, while another participant lauded the feeling of making excerpts, referring to it as "pleasurable." In less controlled contexts, individuals seeing or using LiquidText for the first time routinely praise the putty-like character of the ways objects connect to one another in the workspace. One lesson, then, is that it is possible to make even routine knowledge work tasks more pleasant, and even fun, by providing a different and perhaps more engaging medium. Nonetheless, the ideal way to offer such engagement without distraction from the task at hand is a larger question, and beyond the scope of this research.

CHAPTER VI
SUMMATIVE STUDY METHODOLOGY

With a final, completed version of LiquidText in hand, the conclusion of this research project was a summative evaluation designed to address research questions RQ3, RQ4, and RQ5. That is, I sought to understand LiquidText's impact on the subjective experience, objective outputs, and processes involved in active reading. The evaluation was structured as a controlled, within-subjects, laboratory study, in which participants performed active reading tasks using LiquidText and a more traditional medium. Through several investigative instruments I sought to identify qualitative and quantitative differences between the two active reading events which I analyzed through several means described below. Through the rest of this chapter, I will begin by providing an overview of the study itself, followed by a broad rationale for using this class of evaluation. I will then provide a detailed explanation of the different components and instruments used in the study, as well as background issues like equipment, recruiting, etc.

First, from an overview perspective, this study consisted of several parts. The first was a diary task, followed by the two study conditions (in a counterbalanced order). One condition, the control, involved a single-session active reading task performed using the medium that the participant was most comfortable with—either a traditional PC, or a mix of a PC and paper. The experimental condition included three sessions: 1) learning to use LiquidText and doing a practice active reading activity; 2) performing a second practice activity; and 3) performing an active reading task using LiquidText. Including both conditions, four sessions were required per participant. This is outlined in Figure 6.1.

<table>
<tr><th>Component</th><th>Consists of:</th><th>Serves purpose:</th></tr>
<tr><td>Diary Study</td><td>Participants record instances of active reading.
Runs for 10 days.</td><td>Empirical evidence that participants frequently perform active reading tasks. Allows comparisons based on reading frequency.</td></tr>
<tr><td>Control Condition</td><td>Semi-structured interview about typical reading behavior and reflecting on journals.
Participants perform 55 min active reading task using preferred traditional medium (paper/PC).
Semi-structured interview about task, tools, reading process-differences[2] and difficulties.
TLX, active reading task, and process-difference[2] questionnaires.</td><td>Assess baseline active reading performance.
Determine affect and cognitive load level for task.
Observe details of reading process.</td></tr>
<tr><td rowspan="3">Experimental Condition</td><td>Session 1
Semi-structured interview about typical reading behavior and reflecting on journals.[1]
Explain purpose of LiquidText.
Teach/test all LiquidText gestures.
Participant performs practice active reading task using LiquidText.</td><td rowspan="2">Familiarize participants with LiquidText system.
Help participants learn to incorporate LiquidText into reading process.</td></tr>
<tr><td>Session 2
Brief refresher training/testing on LiquidText gestures.
Participant performs second practice active reading task using LiquidText.</td></tr>
<tr><td>Session 3
Brief refresher training/testing on LiquidText gestures.
Participant performs 1 hour active reading task using LiquidText.
Semi-structured interview about task, tools, process-differences[2], and difficulties.
TLX, active reading task, and process-difference[2] questionnaires.</td><td>Assess reading performance using LiquidText.
Determine affect and cognitive load level.
Observe details of reading process.</td></tr>
</table>

[1] Only performed for first condition. [2] Only performed for second condition.

Figure 6.1. Components of the summative evaluation study.

6.1 Study Design

In creating the summative study for LiquidText, my central goal was addressing my research questions. But within that framework, there were a variety of possible study design options. In this section, I provide a brief explanation for why I chose a comparative, controlled lab study.

6.1.1 Controlled Comparative Study

The first step toward this study design was opting for a comparative study at all, as alternatives were certainly available. For example, I could have deployed LiquidText and conducted a field study of how people use it in the real world, thus ascertaining its impact on their behavior over a longer term and without any sort of control (e.g., [Marshall, Price et al. 2001]).

But there are several reasons why a controlled, comparative study was more appropriate for assessing the ideas behind LiquidText. The first, I would argue, is that any assessment of a new active reading technology should be comparative, because the value it can bring to readers will be relative to what they can already do with their current active reading media. Put another way, LiquidText is not trying to support a wholly new activity, but just support a well-established activity in a better way. As such, its value is going to depend on how it compares to existing technology. Thus, running a carefully controlled comparison just provides a way of keeping this necessarily comparative assessment organized and consistent. Additionally, a comparative, controlled study is consistent with the existing active reading literature [O'Hara and Sellen 1997; Morris, Brush et al. 2007], and likewise most consistent with my research questions which are themselves formulated as comparisons.

But even with a comparative study, the arbiter did not necessarily need to be users. Especially in highly exploratory technology, a theoretical or design critical evaluation can be more appropriate. This is especially true when it is infeasible to provide a complete, consistent, user experience, without which even potentially valuable UI ideas can be misunderstood or overlooked by users. Nonetheless, in the case of LiquidText, a user study seemed most appropriate for several reasons. First, the universal aspects of active reading are well understood on a very broad level (see Chapter 2), but on a fine grained level, are emergent, and highly dependent on the details of the media in use. For example, [O'Hara and Sellen 1997] found PC users tended to adapt their behaviors to the limitations of their medium, taking more notes outside the document itself. Thus to evaluate LiquidText theoretically would be infeasible, since it would not be clear how users would adapt to LiquidText's very different affordances (compared to paper or PCs) and thus how they would behave on a fine grained level. And fortunately, since LiquidText was implemented from an early stage in the design process, it has actually been possible to evaluate its usability over the course of many iterations in order to make it a viable instrument for assessing the ideas about flexible document representations that underlie it.

6.1.2 Lab Study

Beyond just opting for a comparative, controlled study, I also chose to conduct this study in a regulated, laboratory setting. While this is typical of a controlled study, LiquidText could alternately have been evaluated through a controlled deployment, in which participants use LiquidText or other reading media in the course of their day-to-day, real-world behavior, and I observe them doing so. In this case however, such a study would have been infeasible for several reasons.

The first challenge confronting a deployment study of LiquidText is the hardware required to run it. As discussed above, LiquidText requires a sensitive, accurate, reliable multitouch display, and these generally fall into two categories: small and large. And as discussed above, the large devices (i.e., greater than 10" diagonal) tend to have numerous defects. The few displays that can detect over two touch points, which is required for LiquidText, tend to be projected capacitive units that are highly unreliable, and sensitive to seemingly inane factors such as the use of certain florescent lights, being plugged into an electrical outlet, how long the unit has been running, etc. More reliable units are excessively large and complex, and vastly too expensive to deploy. As such, no large hardware platform appears viable for a LiquidText deployment.

Of course, below 10”, there are very usable, reliable, affordable touch sensors, such as in the iPad. But by virtue of using such a small display, these devices would make a comparison study challenging. As I discussed in Chapter 3, an important aspect of active reading is working space—seeing an adequate amount of material at once. If I conducted a deployment study with LiquidText on an iPad, the mere lack of display space would put it at an overwhelming disadvantage compared to paper or to the 24” displays routinely used with PCs.

But even if a viable hardware platform could be found, there are still issues in the LiquidText software itself. As discussed in Chapter 3, active reading is typically performed as part of one’s job or for school, where deadlines and extended work sessions are common. Bringing LiquidText to a level of commercial stability to fit into such an environment would have been extremely difficult, especially given that even low level code for the most basic user interface elements and functions, such as selecting text, had to be written from scratch and would have to be tested and debugged. Moreover, to be used in real world tasks (especially to supplement a typical PC), LiquidText would need to be able to act as a part of a normal document workflow; it would need to be able to open and save to a variety of document formats, and to interact properly with other standard software. Besides the very high cost in time and resources to implement this additional functionality, we then face serious questions of whether LiquidText runs on its own dedicated hardware (and thus forces users to switch hardware platforms repeatedly throughout the day), or else runs on a platform that can accommodate both LiquidText as well as traditional mouse-based applications. Both of these options raise numerous, complex issues and questions that are not central to this research, and thus best deferred to future work. Of course the iPad and other tablet devices are slowly growing a separate touch-based ecosystem of applications, but as discussed above, these tablets are too small for this study.

6.1.3 Within Subjects Design

While between-subjects designs have various merits, the within-subjects design made the most sense for the evaluation of LiquidText for several reasons. The first is practical. To achieve the same *n* for a between-subjects study, I would have had to recruit, and run half the study on, twice as many volunteers as with the within-subjects design. In my past experience with recruitment, including for the formative study, I have seen how challenging it is to find people willing to volunteer multiple-hour lab studies. Thus, I would prefer to get as high an *n* as possible for a given number of participants. Of course, the between-subjects design reduces each volunteer's participation time, but in my experience, marginal changes in participation time do not significantly impact the number of people willing to volunteer. So it is likely I would get approximately the same number of participants in either study design—thus a within-subjects design makes the most sense, as it would result in a far higher *n*.

Additionally, as this study is comparative, a within-subjects design provides another dimension of comparative data. Specifically, participants can report directly on the subjective differences between the two conditions, which can be especially valuable for highly subjective phenomena, like affect. But even for more objective phenomena, this provides a way for participants to report on areas of difference that I may not have known to include in my survey and interview instruments, and thus potentially broaden the differences between the two conditions which my study can identify.

6.1.4 Essay Task

Active reading, unfortunately, cannot be assessed directly. The direct results of an active reading task, generally, are the acquisition of knowledge, development of understanding, and other ultimately invisible phenomena within the reader's own mind. As such, assessing active reading, and with it, the media used in the reading process, requires some ancillary task that draws upon the cognitive structures resulting from the reading process. There are various ways to do this. One can ask the reader to recite passages from the texts they read, testing the participant's memory. Or participants can be asked to explain the gist of the reading, probing more broadly the understanding that was developed. This latter approach is used in two important studies comparing different active reading media ([O'Hara and Sellen 1997; Morris, Brush et al. 2007]). There, participants were asked to write summaries after reading articles, based on the idea that a summary—effectively writing down the gist of the article—can be seen as a canonical active reading task. With my formative study's pilot participants, I followed the same approach but found it inadequate. Part of what separates a summarization task from a recall task is that integration is required, bringing together the various pieces of the reading and identifying the core ideas, themes, etc. But in practice, pilot participants were able to circumvent this by summarizing their reading one section or paragraph at a time, fully relying on their own working memory. As a result, much of the difficulty of the task was eliminated, obviating the need for many of the processes typically associated with active reading (e.g., notes, non-linear navigation, annotation, etc.). As a result, I performed my formative study—as explained above—using a critique task, forcing the reader to reflect upon, and not merely recapitulate, what they read. Generally, I found this to be a more successful route by providing a more challenging active reading task.

6.2 Recruiting

In recruiting participants, my study team (a group of Masters' students who helped me design and set up the gesture refinement and summative studies) and I inquired broadly from a variety of groups. Within the Georgia Tech campus, we posted flyers and posted ads on mailing lists. We also sought a more diverse sample by speaking directly with members of several harder to reach groups, including legal professionals and editors, as these fields are known to involve regular active reading. In practice though, the substantial time commitment of the study (up to 6 hours plus the diary task)

Participant Designation	Background	Active Reading Diary Entries	Gender
P2	Graduate Student, computing	3	M
P3	Undergraduate student, building construction	2	M
P5	Graduate student, business	3	M
P7	Graduate student, computing	7	M
P9	Graduate student, electrical and computer engineering	8	M
P10	Graduate student, digital media	10	F
P11	Graduate Student, electrical and computer engineering	7	F
P13	Graduate student, digital media	8	M
P15	Graduate student, computing	-	M
P17	Undergraduate student, biomedical engineering	7	F
P18	Graduate student, industrial & systems engineering	9	M
P21	Graduate student, computing	9	F

Figure 6.2. Summary of summative study participants.

was more than most full time professionals could offer. The resulting sample then, consisted of a variety of students from the Georgia Tech campus.

Ultimately, twelve participants completed the entire LiquidText summative study (four female), out of a total of 33 who initially signed up.

6.3 Diary Task

After recruiting, the first phase of the summative study was the diary task. The larger goal of this phase was to provide a quantitative means of establishing the regularity with which study participants actually did active reading. I had two reasons for gathering this data. First, as LiquidText is designed to support active reading, it stands to reason that its success in doing so could be dependent on how frequently an individual does active reading—frequent readers, perhaps, might have become more set in their habits, and would have a greater challenge adapting. Thus, in conjunction with other feedback, the diary data could provide a way to explain differences among participants in their use of LiquidText.

Second, the diary data provides a way to justify my experimental method. As I noted above, only when participants are performing the experimental condition of the study are they given practice and training before performing the active reading task.

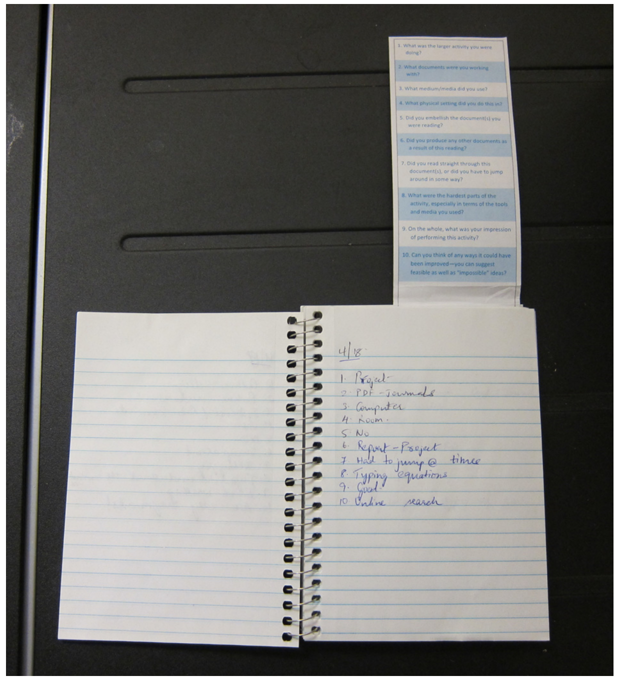


Figure 6.3. A diary used in the diary task from the summative study. Note the bookmark includes the questions participants are to answer with their diary entries.

Those performing the control condition are just given the reading task immediately. This is justifiable if the participants already regularly perform active reading tasks using traditional technologies, since they effectively get "practice" all the time. The diaries thus act as a way to prove this.

6.3.1 Diary Task Methodology

To begin, each participant received a small spiral-bound notebook in which to record their diary entries (Figure 6.3). They were instructed to add an entry each time they completed an active reading task as part of their normal, everyday work; each entry was to consist of answering ten short questions about the reading task (included as Appendix A) pertaining to the difficulties they experienced in the task, what it consisted of, and the like. To help participants remember the questions they needed to answer, a small instruction sheet was included in the pocket of the notebook that gave examples of the types of answers I was looking for. The notebooks also had built-in bookmarks which listed the 10 questions when held up (Figure 6.3).

One of my concerns in designing the diary task was ensuring that participants do indeed regularly record their diary entries, and I took several steps to promote compliance. First, the use of a physical notebook was not incidental; my hope was that the concrete artifact would act as a cue to remind participants to add entries. Second, I sought to minimize the perceived burden of recording entries by emphasizing that each diary entry can be quite simple, with only a few words per question. Finally, I emailed participants each day to ask how many entries they had added so far; the goal was to remind them to reflect on their reading that day and record any forgotten entries.

The diary task was run for a total of ten days per participant, and was completed before participants could move on to the following phases of the study. To note, 11 of the 12 participants returned their notebook (see Figure 6.2), and two recorded entries past the 10 days I required.

6.4 Opening Interview

As noted above, the summative study consisted of four sessions: three for the experimental (LiquidText) condition, and one for the control. Of course these were counterbalanced to compensate for learning effects that could carry over from one condition to the other. But whichever condition came first, the initial session always began with a ten minute, semi-structured opening interview.

The purpose of this opening interview was to enquire about participants' typical active reading behavior and their general background. Specifically, I asked participants about three topics. First, I asked them to reflect on the diary task and any general thoughts they wanted to volunteer about their reading behavior. Second, I enquired about their background, including a discussion of their educational status and what touch screen devices they use, and a discussion of the types of reading they typically perform. Third, I asked participants about the details of their active reading behavior in terms of a) embellishing, b) creating external notes/outlines, and c) performing comparisons. As with the diaries, I gathered this data was because it offered the potential to aid in explaining differences in participants' use of LiquidText which I might have observed later in the study.

6.5 Control Condition

As described above, the summative study's control condition was meant to establish a baseline for each participant's performance of an active reading task. That is, to find out how they do active reading using normal, typical technologies. From that perspective, I had many options; I could have compared LiquidText to media such as tablet PC's, eBook readers, desktop PCs, textbooks, etc. But confounding this flexibility is that participants are not *trained* in the control condition. This again is acceptable because they already perform active reading using traditional technologies on a regular basis—but it also means that the technologies used in the control condition must indeed be the ones with which participants regularly work and with which they feel comfortable. However, as I observed in my formative study, people are often habituated to a particular reading medium—such as paper or PCs. Therefore, I could not simply force all participants to use the same reading medium for the control since, in many cases, participants could well be using a medium with which they have little experience in active reading. Therefore I designed the control condition to adapt to the participant, offering a choice of medium, so that each participant is assigned a medium with which they are comfortable.

But while I needed the control condition to adapt, this adaptation could not be boundless. It was important to maintain some level of consistency within the control condition so that, apart from the flexible, touch-based document representation, the control condition and experimental condition would be as similar as feasible. For example, I sought to provide comparable amounts of display space and comparable keyboards, in both conditions. Thus, I limited the control condition to two technology

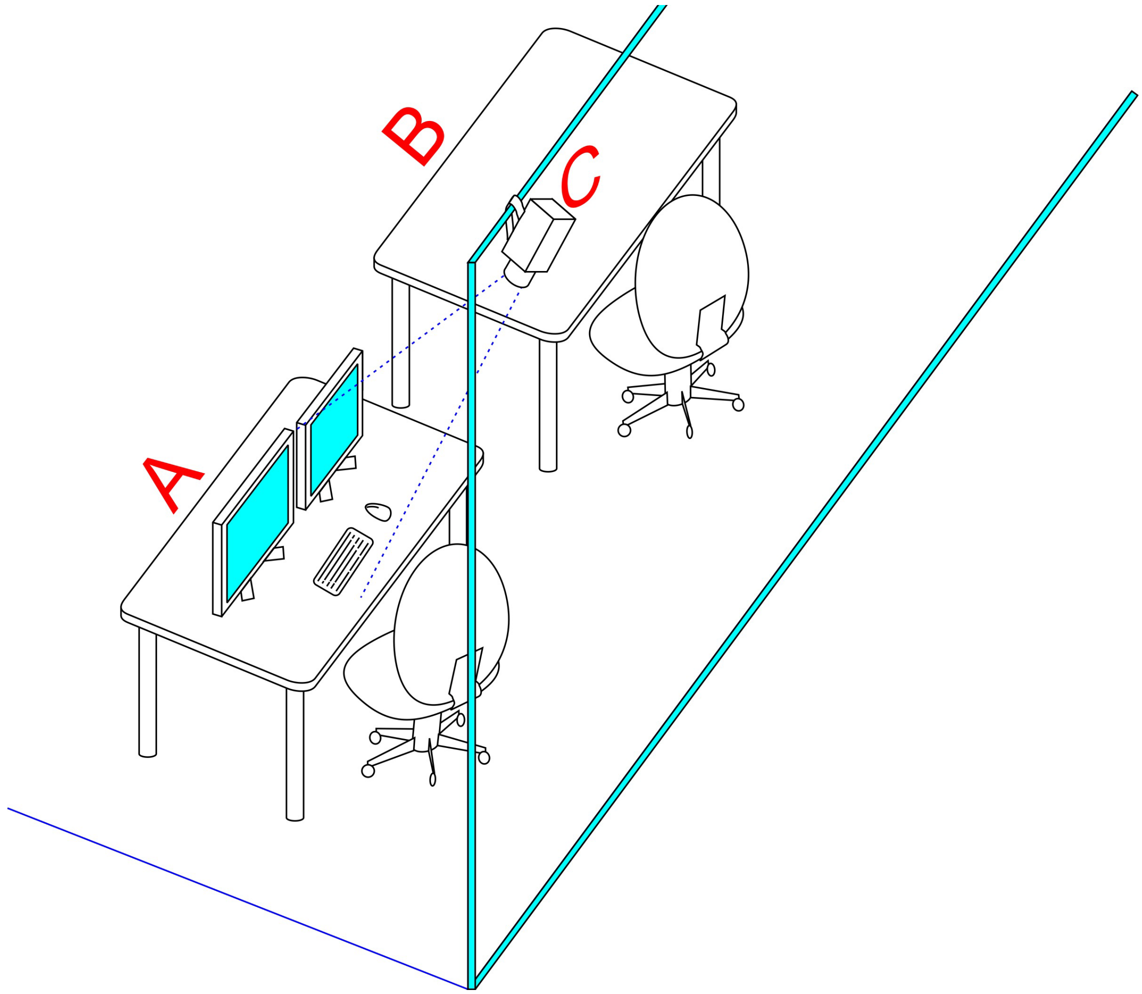


Figure 6.4. The PC environment of the control condition. Participant sits at A, experimenter sits at B. Participant is recorded by camera C. (Orthographic projection, not exactly to scale.)

configurations: a purely PC-based reading environment, and a PC/paper hybrid reading environment. I spoke with participants at the start of the control condition and enquired about their typical reading technology, and assigned them to the environment most closely matched to that with which they considered themselves most comfortable.

6.5.1 The PC Environment of the Control Condition

For both conditions, the study took place in a large lab with two desks in proximity: one for the participant and one for myself, as I was the experimenter. This enabled me to easily conduct the interviews and answer any participant questions during the active reading task. In the PC environment, the participant's desk had a dual-monitor

PC running Windows XP, and a small, wireless keyboard and mouse (Figure 6.4 A). Both desks were 48" x 28.5, and the participant's monitors included a 20" unit with 1680x1050 resolution, and a 19" unit with 1280x1024 resolution.

To perform the task, the participant's computer had Microsoft Word 2007 installed. After being assigned to the PC environment, I asked the participant whether they were familiar with Microsoft Word 2007. All participants responded that they were familiar with it, though to be sure, I opened a Word window on the participant's PC and briefly demonstrated how to highlight and create comments in a Word document. Next, I opened the article that the participant had to read, and an empty Word window in which to write their response. The article was formatted in Times New Roman, 12 point font, with 1.15 x line spacing. I left the windows positioned arbitrarily, but instructed the participant to position the windows as desired across the two monitors.

6.5.2 The PC + Paper Environment of the Control Condition

The basic room layout for the PC + paper environment was the same as in the PC environment, but the materials given to the participant were quite different. First, rather than receiving the article they were to read as a Microsoft Word document, they received a paper printout. Second, they received several sheets of blank notebook paper for taking notes during their active reading task, as well as a pen, a pencil, and a collection of colored highlighters for annotation. Participants still used the computer, but this time only for writing their responses. But as participants here had separate sheets of scrap paper, and a separately printed copy of the article which they could read and annotate, I shut off the second, smaller monitor to try to keep the working space between the two environments comparable (see Figure 6.5).

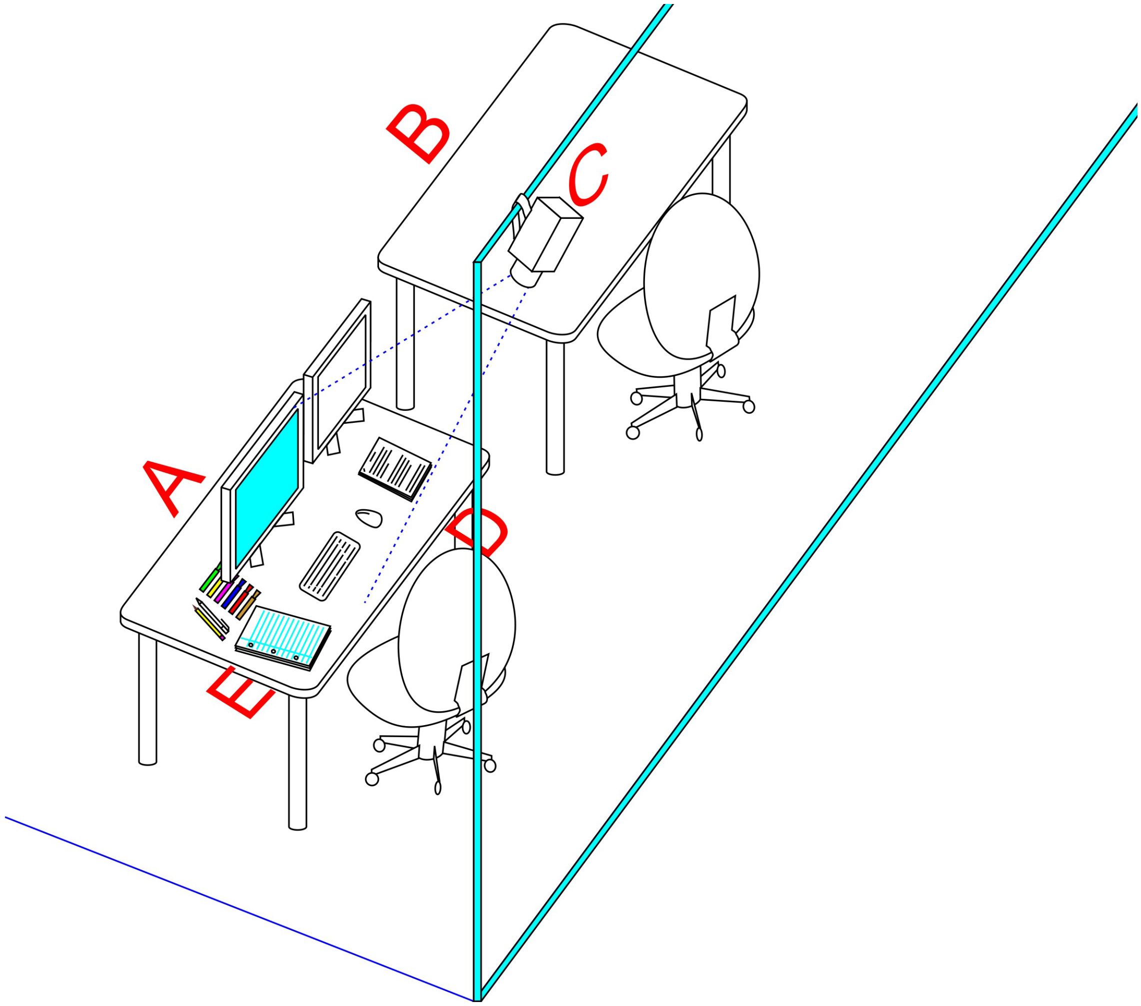


Figure 6.5. The PC + Paper environment of the control condition. Participant sits at A, experimenter sits at B. Participant is recorded by camera C, and receives materials D and E. (Orthographic projection, not exactly to scale.)

As in the PC environment, I proceeded to introduce the participant to the materials they had before them. I first gave the participant the notebook paper, pen, pencil, and highlighters (6.3 E), and also made sure they were comfortable using Microsoft Word. I then gave them the printed article (Figure 6.5 D) to read and opened an empty Word document for them to write their response. To note, the 4-page article was stapled in the upper-left corner, and printed single-side in Times New Roman, 12 point font with 1.15 x line spacing.

6.5.3 The Reading Task

After conducting the opening interview, if it was the first condition for that participant, and introducing the materials available for completing the active reading task, I finally gave the participant the reading task itself. As discussed above, the reading task required the participant to read an article and write a critique about it, all in 55 minutes.

But before giving participants the article, I first described the task. I explained that in their critique, they were to focus on thinking deeply and analytically about the article, and that the mechanics of their writing were of lesser importance. I told them that they could even write a detailed outline if that would be better. I also generally warned them that 55 minutes is a short time, and that it would be important for them to balance their reading and writing time. Finally, I provided them with a printed prompt explaining their task in more detail (see Appendix B). I offered participants a few-minute break to prepare for the task, allowing them to arrange the materials on their desk, adjust their chair, reflect on the task, etc. Once they were ready, I then provided the article, either opening the document on their computer or giving them the printout.

I gathered several types of data through the critique task. First, participants were video recorded from above-behind (Figure 6.4 C or 6.3 C). Second, I saved participants' annotated copies of the articles they read, whether in electronic or paper form. Finally, in the PC + paper environment, I saved participants' note sheets; only one participant made a note document in the PC environment but deleted it before I could save it.

The Articles

To understand the nature of this critique task, it is important to consider the articles that participants received. Since this is a within-subjects study, I had two different articles which, along with condition order, were counterbalanced using a Latin Square method. In this section, I discuss several aspects of these pieces in more detail.

First, in choosing articles for this task, my larger goal was to give the participant an active reading experience that was realistic and representative of many real world tasks, while being brief enough to fit in 55 minutes. Thus, I had several requirements for the articles used in the study: 1) The article needed to be difficult enough to require *active* reading. Part of this was sheer length—I did not want the entire article to fit readily in the reader's working memory, or there might be little need to externalize the one's thoughts in the form of notes or annotations. Comparisons would likewise become trivial as well. 2) Besides length, I wanted the article to be complex. Even if the document was adequately long, if it merely went into detail about concepts that were simple, or few in number, it could still be easy for the participant to manipulate the article fully within their mind. 3) Conversely, it was likewise important that the article not be too long or too complex. As happened in my formative study, an excessively long, dense document requires so much time to read, that participants may have very little time left to reflect on or analyze it. 4) But even if the length and complexity are appropriate, the subject matters as well. Since the task is a critique, it is important that participants can somehow relate to the article—i.e., that they can form an opinion about it. Thus, the popular science articles used in other studies (e.g., [O'Hara and Sellen 1997]) would likely not be adequate. 5) Finally, even though I used counterbalancing, I wanted to keep the both of the articles similar in difficulty.

To find such articles, my study team and I searched the web for 4-7 page pieces that included controversial opinions, or some form of argument, that could be critiqued. We considered how much time we required as we read the pieces ourselves, and also observed how long our initial study participants spent reading the practice articles used in the LiquidText condition's training sessions. Based on our observations, we focused on approximately 4-page articles, and ultimately selected two popular pieces from the New York Times website that discussed philosophy and religion; the pieces are 2034 and 1999 words respectively.

6.5.4 The Post-Task Questionnaires

Immediately after completing the critique task, I administered several survey instruments to obtain broad and structured feedback from the participant. This included

1) a NASA TLX cognitive load questionnaire, 2) a general active reading task questionnaire, 3) a medium-specific (i.e., specific to the medium) active reading questionnaire that varied between the PC and PC + paper environments, and 4) a questionnaire comparing the two conditions, which was of course only administered during the second condition. I discuss each of these in turn.

The first questionnaire that I administered was the NASA TLX. I gave this to participants immediately after they completed the critique task, in an effort to gauge their cognitive load while the experience of performing the task was still fresh in their minds. The reason for measuring cognitive load is because it can 1) provide additional data on the subjective experience of using LiquidText compared to other media, and is thus relevant to RQ3, and 2) it had the potential to help explain other findings by showing the role of cognitive load how LiquidText is used or received.

Next, I administered the general active reading task questionnaire. This 24 question, Likert-scale (1-5) survey asked a range of questions about the participant's subjective reactions and the processes involved in the reading task. It enquired about both the importance and the difficulty encountered in various aspects of the reading process, such as comparison, note taking, content arrangement, embellishment, etc. The purpose of this questionnaire was directly related to RQ3 and RQ4, to understand the differences in subjective experiences and reading process associated with the different conditions. Since this questionnaire is administered in all environments and conditions, it readily supports such comparisons.

After the general questionnaire, I administered one that is far more specific to the particular technology the participant used. As such, this questionnaire varies not only between the control and experimental conditions, but between the PC and PC + paper environments within the control condition. This questionnaire also uses 5 point Likert scales, but focuses less on supporting comparisons. Rather this questionnaire is for providing more insights into the nuances of performing the task with a given medium, and as such feedback can potentially help explain other findings.

The final questionnaire asks the participant to directly compare the experimental and control conditions, and thus is only run after whichever of the two conditions is administered last. This questionnaire asked the participant to reflect on several parts of their experience, including 1) their overall preferences between the media used in the two conditions, 2) which medium was more appropriate for certain aspects of the task or under certain circumstances, and 3) which medium seemed to lead the participant to better performance on the task. This questionnaire touches on RQ3, RQ4, and RQ5, as it enquires into differences in subjective experience, process, and even the outputs of the two conditions.

6.5.5 Post-Task Interview

While the questionnaires were designed to provide a broad, consistent, structured perspective on participants' experiences, the semi-structured, post-task interview allowed the freedom to go deeper into issues particularly relevant to a participant. The major discussion points of these interviews closely reflected the topics of the above questionnaires, and so included several sections: 1) a general post-task interview, 2) a medium-specific post-task interview, and 3) a post-task interview comparing the media. Since these discussion points are very similar to those of the questionnaires, I refer the reader to the previous section for more detail.

6.6 Experimental Condition

While the control condition sought to establish a form of baseline for participants' processes, experiences, and outputs in active reading, the experimental condition provided a way to see how these things are impacted by LiquidText. And in contrast to the single-session control condition, the experimental condition required three separate sessions. The first included the opening interview, teaching the participant how to use

LiquidText, and providing practice using the system. The second mainly included more extensive practice using LiquidText. And the third session included performing the actual active reading task, followed by approximately the same questionnaires and interviews as in the control condition. These three sessions were always on separate days; I sought to keep them within a two-week window, though on some occasions that was not possible.

The protracted nature of the experimental condition comes from our need to teach LiquidText to participants. The control condition is predicated on participants' prior familiarity using the given medium, but with LiquidText, the participant is at a deficit in two ways: 1) they are not familiar with the basic user interface widgets and conventions of the LiquidText UI, and 2) they are not experienced with many of the particular tools that this unfamiliar interface offers. So even though LiquidText supports the same general active reading processes as other media, it does so in an extremely different manner. As such, the experimental condition is designed both to familiarize the participant with LiquidText's tools and functions, but moreover to allow the participant time to figure out how to assimilate those tools and functions meaningfully into their active reading process. In other words, I needed to make sure that participants knew more than just how to perform a function, but that they knew when and where performing that function would make the most sense. Several factors went into promoting this, including a scenario-driven demonstration LiquidText's functionality, as well as over 70 minutes of practice performing active reading tasks with the system, all of which are described in detail below.

Throughout the rest of this section, I explain the structure of the experimental condition. To begin, I describe the environment and equipment with which the study took place, and then discuss each of the three sessions in turn.

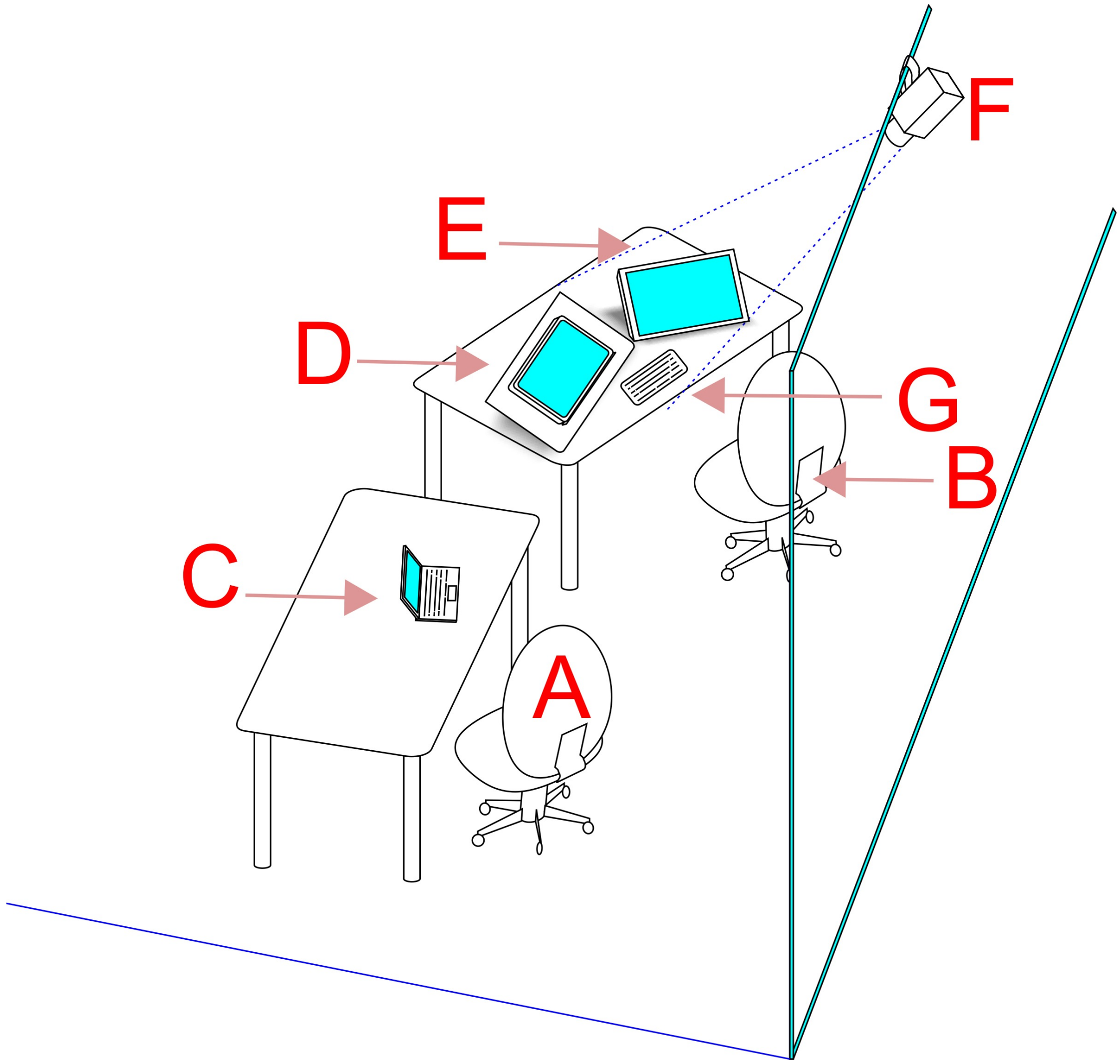


Figure 6.6. Study environment for experimental condition. Experimenter sits at A and participant sits at B. Experimenter can explain LiquidText using small touch-screen computer C. Participant uses dual-touchscreens D and E, and keyboard G.

6.6.1 Environment and Equipment

The testing lab for the experimental condition was a modified form of that used in the control (Figure 6.6). This was advantageous in allowing me to switch the lab from one setup to the other between participants. As in the control condition, there were two desks in close proximity: one for the participant, and one for the experimenter. But unlike in the control condition, the experimenter's desk now included a small multitouch tablet (Figure 6.6 C) which was used for demonstrating LiquidText's interactions to the

Figure 6.7. Modified multitouch laptop used as the primary monitor for the experimental condition of the summative study (parts of hinges were still visible when study was run).

participant. Likewise, the desks were significantly angled so the participant could actually see the tablet's display.

The participant's desk setup was also very different from in the control condition, now containing two angled, multitouch displays. I discuss each in turn. The first unit (Figure 6.6 D) was the primary display and ran LiquidText. It included a projected capacitive multitouch sensor as described above, but was also a complete laptop (a version of the Dell Studio 17) since this form of touchscreen could not feasibly be purchased as a separate item. To make this laptop into a tablet-like form factor, I disconnected the hinges and rearranged some of the wiring in order to flip the display backward over the bottom of the device. The unit was then positioned upside down (so the power cord could be plugged in without getting in the user's way) on a mount I made atop a Cintiq stand (Figure 6.7).

By being built on a Cintiq stand, the mount allows easy and reliable angle adjustment. This is important because the user faces a tradeoff between a shallower angle (compared to the desk) that will be less tiring for their hands [Sears 1991], versus a higher angle that will cause the screen to appear discolored (since the laptop used a typical twisted nematic display designed to be viewed close to the normal). Although I adjusted the display's contrast curve to partially compensate for the distortion, this had limited effect, and so the mount enabled participants to find the best tradeoff for their particular preferences. A second purpose of the mount is to provide a place for the participant to rest their hands as they use the touchscreen, since a hand rested on the display itself can easily be mistaken for an input point. Therefore, the mount includes several inches of room on either side of the tablet (Figure 6.7).

The second display on the participant's desk (Figure 6.6 E) showed the word processor, and was connected to the converted tablet as a secondary monitor. In contrast to the 4-finger detection capabilities of the tablet's projected capacitive touch sensor, this second monitor used simple optical detection that allowed it to reliably sense only one or two fingers, which was adequate for the word processor.

In front of the multitouch displays was the keyboard, which was used for text entry in writing comments as well as the response essay (Figure 6.6 G). Two notable features of the keyboard were its size and that it was wireless. The latter was important because I expected users to change between annotation and response writing repeatedly, and thus switch their focus between the two displays. As such, I did not want anything, such as a wire, to inhibit their repositioning the keyboard. The former issue was that the keyboard was unusually narrow as well. This allowed the keyboard to be left in front of the first multitouch display without preventing the user from resting their elbows on the table while they performed touch interactions.

The final significant environmental concern was glare. As seen in Figure 6.7, the multitouch displays had glossy screens and easily reflected any well lit items in the room. One approach to mitigating this would have been to dim the room lights for the experimental condition versus the control, however some research has found significant effects of ambient lighting [Knez 1995] on cognition, so it was important to keep the lighting consistent across conditions. Thus, I found an area of the room with solid-colored walls, so there would be no specific features to reflect. I also disabled the room lighting immediately above and behind the desks to minimize even the solid-color reflections, though I kept the lights in front of the desks activated so the room would feel well lit to the participants.

6.6.2 First Session of the Experimental Condition

As noted above, the experimental condition included three separate sessions, the first of which concerned teaching the participant to use LiquidText, and providing practice time for the participant to decide how best to appropriate LiquidText's functionality. Thus first session began (after conducting the opening interview, if necessary), with teaching the individual LiquidText gestures.

I began the gesture tutoring by introducing the participant to the LiquidText screen layout, the two monitors, and briefly explaining the purpose LiquidText is meant to serve. Then, for the 24 primary LiquidText functions, I explained how the function might be used, and demonstrated it on the multitouch laptop on my desk (Figure 6.6 C). After each gesture, I asked the participant to try to repeat what I had demonstrated. The tutoring required approximately 20 minutes.

The next phase of session-one was to contextualize the gestures—to show the participant how they *could* be used in an actual task. The need for this came out of the LiquidText formative study, where participants explicitly asked for an idea of how they should use LiquidText's functionality—i.e., some best practices for this unfamiliar system. Thus, I created a simple task modeled on how one might use LiquidText to study for a test. For each step of the task, I introduced a need (e.g., "We've come to an important date that we'll need to know...") and concluded with how one of LiquidText's features could be used to address it (e.g., "…so we highlight it green."). Since the participant had just learned the LiquidText functions, I performed each step of the task on my touchscreen laptop before asking the participant to do the same on their system. This phase also required approximately 20 minutes per participant.

The third phase of session-one was a 25 minute practice active reading task; since it was only a practice, the participant was asked write a summary of an article, rather than a critique. I began this phase by giving the participant a laminated Gesture Reference Sheet, which contained graphical depictions of the LiquidText gestures I thought would be most easily forgotten. The participant then received a copy of the task instructions and had up to 3 minutes to adjust their workspace, review the instructions, etc. Finally, I opened a practice article in LiquidText on the participant's system, and gave them 25 minutes to write a summary, emphasizing that this was only a practice, and meant as an opportunity for the participant to figure out how best to use LiquidText's features. I also instructed them to ask me if they had any questions, forgot a gesture, etc.

After completing the summarization task, the next phase of the first session was a Gesture Fluency Test. Here, I simply went through the 19 LiquidText functions I considered most important, and asked the participant how they would perform each one. The purpose of this phase was partially for me to understand how well the participant had mastered LiquidText's gesture interface thus far, but also as a way to reinforce those gestures, since working to recall information tends to support its memorization [Roediger and Karpicke 2006].

Finally, I dismissed the participant, and instructed them to take the Gesture Reference Sheet. I asked that they briefly review the Sheet daily if possible to promote retention of the gestures.

6.6.3 Second Session of the Experimental Condition

The goals of the second session of the experimental condition were very similar to those of the first, but with more of a focus on function appropriation than on tutoring.

Still, the session began with a brief refresher of LiquidText's functionality. Here, I went through the same 24 functions as in the tutorial, and asked the participant to perform each one. If they showed uncertainty, I gave them the answer and showed them how to perform the function correctly. There were three objectives here: 1) to ensure the participant knew *what* functions LiquidText supported; 2) to refresh the participant's memory of *how* to perform the gestures; and 3) to support longer term memorization through recall. This phase required about 20 minutes.

After that, the second session was very much like the first. Its second phase was exactly the same as the practice task from the first session, except that the participant had a different article and 45 minutes to perform the task. The third and final phase was a repeat of the Gesture Fluency Test administered in the first session, and performed for the same reasons.

6.6.4 Third Session of the Experimental Condition

The third and final session of the experimental condition was very similar to the control condition, and is where participants were given the actual, graded active reading task using LiquidText.

As in the second session, I again started by providing a refresher of 24 of LiquidText's functions. And also like the second session, I then administered an active reading task. This time however, the reading task was structured the same as that of the control condition, and so was 55 minutes in duration, and required the participant to write a critique in response to the article they received. During this task, the participant still had the Gesture Reference Sheet.

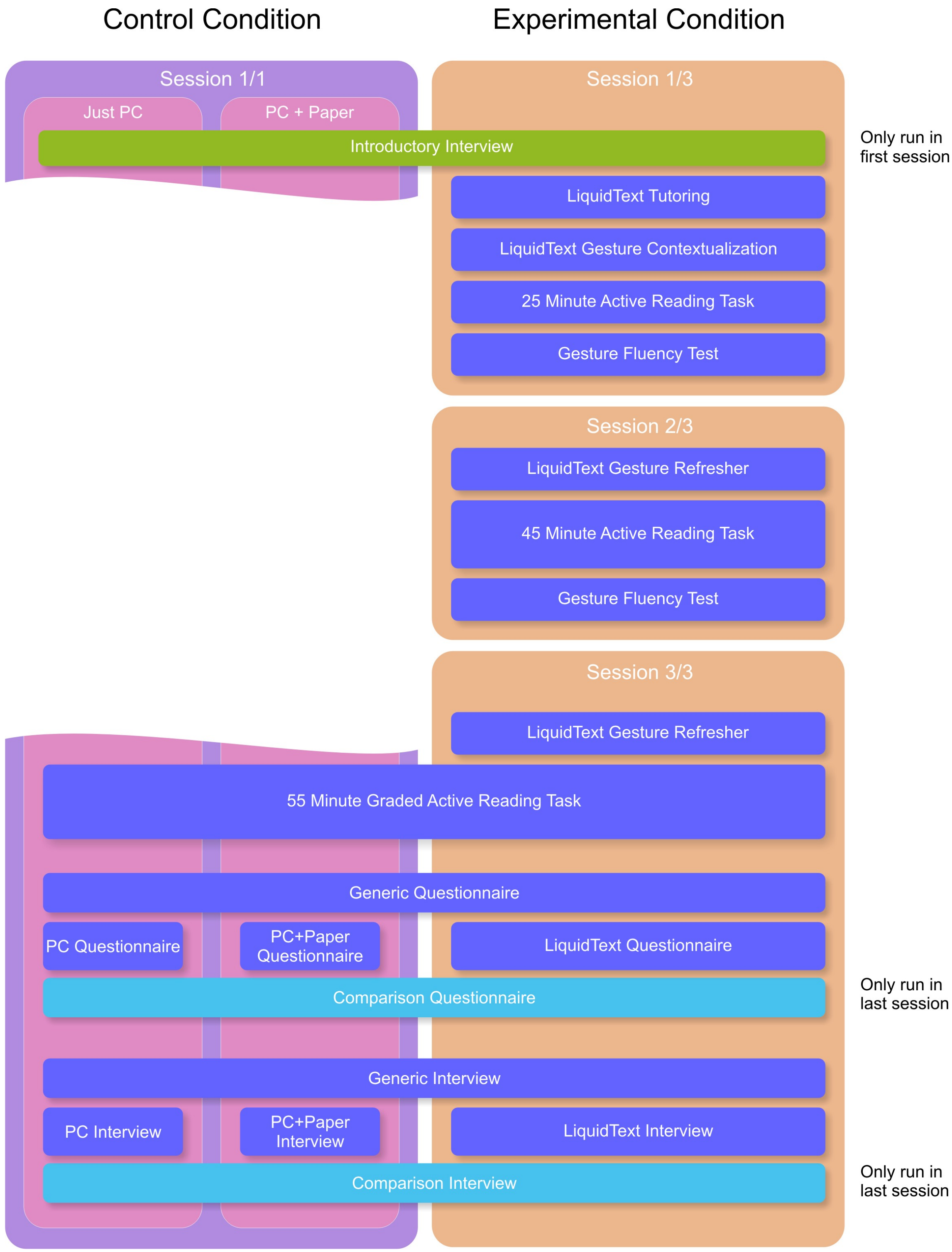


Figure 6.8. Graphical depiction of where the various data gathering instruments, tasks, and tutorials were administered throughout the laboratory portion of the study, especially what was common among different conditions, environments, and sessions.

Data Source	P2	P3	P5	P7	P9	P10	P11	P13	P15	P17	P18	P21
Diaries									x			
Opening Interview	x		x	x	x					x		
Experimental Condition Post-Task Interview	x		x		x	x	x					
Control Condition Post-Task Interview	x		x	x	x			x	x			x
Experimental Condition Response Essay												
Control Condition Response Essay												
Experimental Condition Task Video			x									
Control Condition Task Video			x									
Questionnaires												

Figure 6.9. Areas where data was saved (green) or lost (red, x) for each participant.

Upon completing the active reading task, the experimental condition proceeded exactly as in the control condition. First, the participant completed several surveys: 1) a NASA TLX cognitive load survey, 2) a generic post-task questionnaire, 3) a LiquidText-specific post-task questionnaire, and 4) a comparative survey, if I was conducting the last condition for the participant. Second, I conducted the semi-structured interviews including: 1) a generic post-task interview, 2) a LiquidText-specific post-task interview, and finally 3) a comparative interview if I was conducting the last condition. As in the control condition, the interviews and questionnaires took approximately 30 minutes.

A summary of the different components that were administered during the different parts of the summative study is shown in Figure 6.8. Note the wide boxes indicating instruments used in multiple parts of the study.

6.7 Data Gathered

As discussed above, I used several means to capture data throughout the study, including diaries, video-recording, audio recording, and questionnaires. But while participants were generally comfortable being recorded and providing data, I experienced considerable challenges with my data gathering equipment. While I tested my audio/video recording equipment at several points during the study, my tests were not frequent enough (or long enough) to discover that silent, intermittent errors were occurring in the audio system (and far less frequently, the video system). As a result, a portion of my interview data was irretrievably lost (see Figure 6.9), and several gaps occurred in the videos of participants performing the tasks. Most of the data though, were properly saved and analyzed. In the next chapter, I summarize this analysis process, and explain in detail the findings from the formative study.

CHAPTER VII
SUMMATIVE STUDY RESULTS

As explained in chapter 6, the purpose of the summative study was to understand the impact of a flexible, high degree-of-freedom representation on active reading. In this chapter, I present the findings of this study in the context of my latter three research questions. Again, these are:

RQ3: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the <u>subjective experience</u> of active reading?

RQ4: What is the impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the <u>process</u> of active reading? To what extent do users appropriate the affordances of the system, and to what extent do reader processes change in response?

RQ5: What is the impact of a computer mediated, highly flexible, high degree-of freedom text representation on the <u>results and outputs</u> of active reading, vis-à-vis more conventional media?

I begin this discussion with an explanation of the analytical methods I used to derive my findings, followed by a discussion of the findings themselves in terms of RQ3, RQ4, and, finally, RQ5.

7.1 Analysis Methods

In order to make sense of the various types of data I gathered, I employed several different analytical approaches. Here, I discuss each of these, and explain why I chose each approach.

7.1.1 Diaries

As I explained above, participants used the diaries to record certain facts about their active reading experience prior to the start of the lab portion of the study. In contrast to the complex diary analysis described in Chapter 3, the goal of this analysis was mainly to gather statistics on participants' reading habits in order to identify correlations with the reading behaviors shown during the lab study. To do so, I first identified the diary entries that fit the definition of active reading I have been using in this research. This includes entries involving non-sequential navigation, annotation, the production of output documents or notes, the use of multiple documents at once, or content searching within or among documents. Within this group, I tallied diary entries that include specific attributes, such as whether documents were embellished, whether output documents were produced, whether there was non-linear navigation, and what reading medium was employed.

As a result of this simple analysis, the diary data were brought into a quantitative form, where they could be studied with the various statistical techniques described throughout this chapter.

7.1.2 Interviews

In order to analyze the interviews, I sought a relatively general approach guided in part by the type of data required to answer my research questions. First though, after having the interview recordings transcribed, I chunked and labeled the transcripts. My unit of analysis was roughly such that each chunk could be reduced to one fact about the participant or their beliefs, such as "Highlighting is easier on paper," or "LiquidText is a more unified way to make excerpts—all just in one app." This resulted in 573 chunks.

In analyzing these chunks, I performed an open-coding analysis, organizing the labels into categories. Generally, I let categories emerge from the labels themselves, but given the larger goal of this study, I also created some categories based on my research questions, including the reading processes used with each medium, what people liked/disliked about each medium, etc. This resulted in 23 top-level categories.

To further understand participants' reading processes, I analyzed them as sequences. This entailed breaking participants' descriptions down into simple, sequential outlines, showing which general steps they performed in the active reading task. These included things such as creating a reading strategy in advance, highlighting the article, or making multiple passes. Putting participants' descriptions into this common format and vernacular, I was better able to compare their processes and look for patterns among them.

In total, these methods of analysis allowed me to readily identify how participants shed light on my research questions through their interview responses. This was qualitative, in the structure of the categories that emerged, as well as quantitative, in the densities of each category.

7.1.3 Questionnaires

In contrast to the interviews, I did not need to perform an analysis to reduce my raw data into a manageable qualitative or quantitative form. The questionnaire data were rather entered directly into spreadsheets where they could be investigated with statistical techniques described throughout this chapter.

7.1.4 Essays

As the sole instrument for assessing participants' active reading performance, the response essays served a critical purpose in this summative study. To assess them—and, by proxy, active reading performance—first required a rubric.

In developing a rubric, there were several important factors, including differentiation. That is, I needed a rubric that would clearly identify many areas where the essays differed, but comparatively ignore areas where they would all be similar. To do this, I required a rubric tailored to the particular type of essays participants wrote—critiques of an argumentative article. But as there is no standard, widely accepted rubric for such an assignment, I worked with a literature professor to develop one specifically for this research.

To develop this rubric, I began with a very general writing and communication rubric from the Writing and Communications Program at Georgia Tech (see Appendix C), which itself was based on a validated, widely accepted series of rubrics from Iowa State University. I then modified this rubric, adding and removing dimensions of assessment to better align it with the critique task, especially as encapsulated in the prompt given to participants. The resulting rubric had seven dimensions of assessment, each of which ranged from a score of 1 for “basic” competence, to 6 for “exemplary.”

I assessed this rubric using a formative grading workshop, with the primary goal of ensuring different graders could use it to arrive at similar scores. For this, I invited three Interactive Computing graduate students who were not otherwise involved in the study to grade two sample critique essays (these were from participants who dropped out of the study before completing both conditions). Using the draft rubric, the graders independently assessed each essay and then discussed their respective scores, focusing on areas where the scores differed between graders or where the rubric was ambiguous. The workshop found that the draft rubric largely worked as intended, but also led to various changes in the rubric’s language to resolve ambiguities and potentially lead to more consistent scoring. This revised, final rubric is shown in Appendix D.

With the rubric completed, I recruited two people not otherwise involved in the study to act as essay graders. I initially sought students from Georgia Tech’s School of Literature, Communications and Culture, but making the necessary arrangements proved infeasible. Instead, I selected graders from the School of Interactive Computing, including a post-doctoral researcher and a senior doctoral student studying Human Centered Computing. Following the methodology used in the formative grading workshop, the graders first independently scored the essays. After each one, they attempted to reconcile their assessments, discussing any differences between their scores and having an opportunity to revise them. The resulting reconciled and un-reconciled scores then served as input for various subsequent statistical methods used to explore LiquidText's impact on active reading.

7.1.5 Active Reading Video Data

The final type of data included in my analysis was the video recordings of participants performing the active reading tasks. The objective of this analysis was to help elucidate participants' reading processes by reducing them to a form where they could be sought for patterns, or where quantitative methods could be used to identify statistical relationships.

Thus, the analysis began with video coding. The codes were relatively simple, and corresponded to the major tasks performed in active reading, as well as more specialized capabilities available only in certain media. This included events such as moving through the text and adding comments, as well as functions such as collapsing the text in LiquidText. The coding was performed by four students whom I recruited and trained in the use of the codes. The final code sequences were brought from our video analysis application, ELAN[6], into Microsoft Excel where they could be visualized and analyzed quantitatively.

7.1.6 Analysis Notes

Throughout the rest of this chapter I discuss results obtained through several statistical methods. Unless otherwise noted, the threshold for statistical significance used in these cases is $\alpha = 0.05$. Additionally, references to ANOVA refer to single factor Analysis of Variance tests, and correlation coefficients refer to Pearson's r.

---

[6] Available at http://www.lat-mpi.eu/tools/elan as of 12/22/2011.

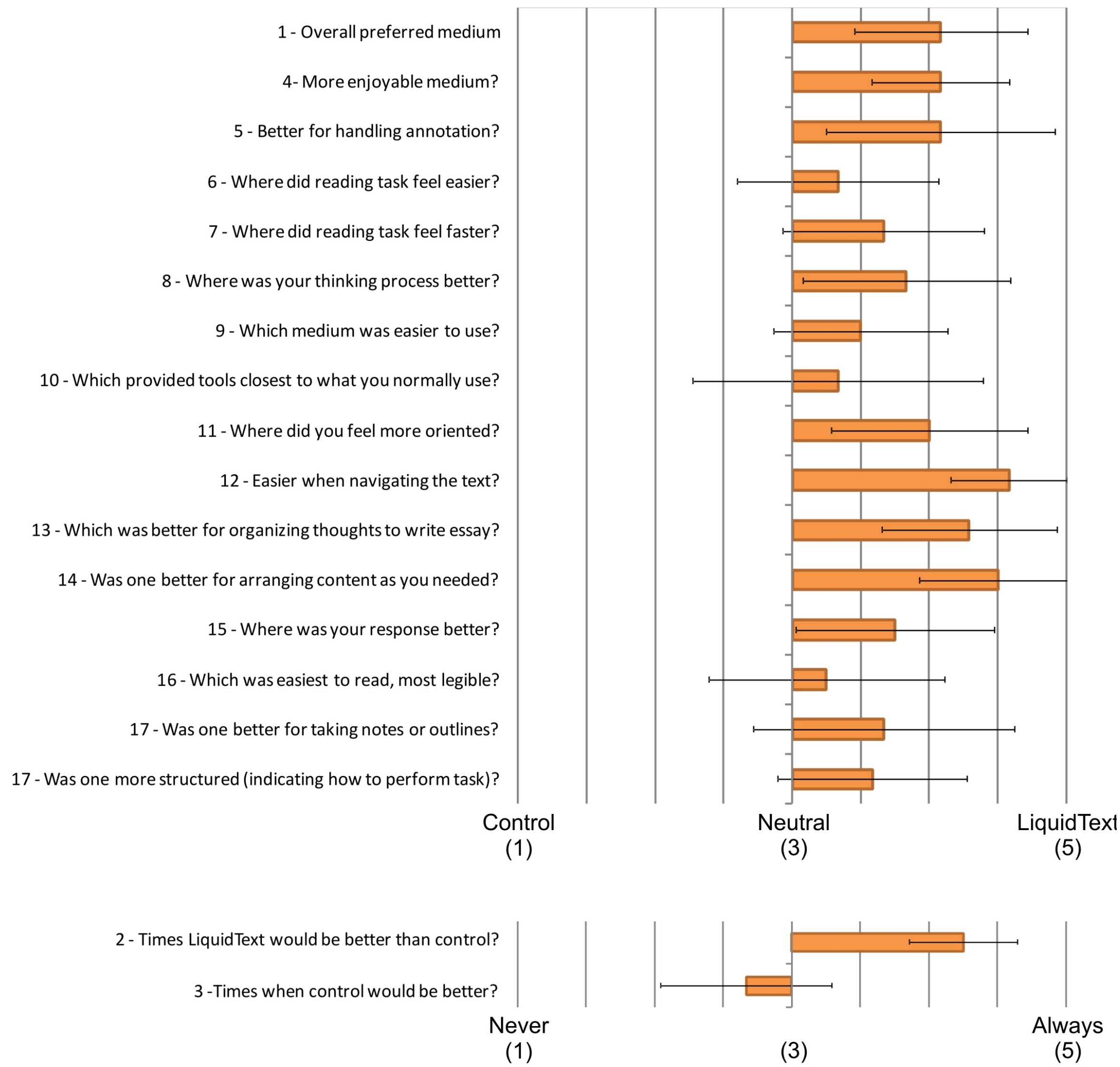


Figure 7.1. Mean participant responses to Likert-style questions comparing LiquidText to the control medium. For the first group of questions, a response of (1) corresponded to the control medium, and (5) corresponded to LiquidText. For the second group, (1) corresponded to "never" and (5) corresponded to "always". Error bars represent a 95% confidence interval. Questions are reworded here for brevity.

7.2 Subjective Experience

To understand the impact of LiquidText's representation on the experience of reading, I enquired of participants in several ways. First, I asked them to make direct, retrospective comparisons, recalling how one medium compared to the other. I also made indirect comparisons by asking a variety of experiential questions independently for each condition, which I then compared in the analysis. In this section, I report on these findings, providing: 1) a broad overview of participants' reactions to LiquidText as an active reading medium, 2) a detailed discussion of the positive impacts of LiquidText on participant experience, 3) a detailed discussion of the negative impacts of LiquidText on experience, and 4) a concluding, overall picture of this technology's impact on experience.

7.2.1 Broad Assessments

To begin, I provide a broad picture of participants' reactions to LiquidText. Figure 7.1 shows mean participant responses to Likert-style, direct comparison questions, and highlights the fact that in virtually all areas (neglecting margin of error) LiquidText was preferred. Question 1, for example, asked participants to indicate which medium they preferred—1 for the control and 5 for LiquidText—and the sample mean response was 4.08, a strong preference for LiquidText (and so, with 95% confidence, the population mean would be *at least* 3.45, and so still a preference for LiquidText). Likewise, in Figure 7.1 Q2, participants felt LiquidText would, far more often than the control, be the right medium to use for a given task. Critical to answering research question three, though, is to qualify and elaborate upon this general preference. In that vein, there appear to have been several factors that contributed to participants' general preference for LiquidText.

One of these factors is enjoyment. Question 4, from Figure 7.1, shows participants consistently found LiquidText to be significantly more enjoyable than the control medium (mean answer on Likert-style question was 4.08/5).The interviews echoed this apparently broad, positive affect. In reference to the system, P18 said, "I love it, actually." And P17 described how LiquidText would impact typical student tasks,

> *"…and that is stuff that students hate to do, where [LiquidText] kind of made it fun because it's like I get to play while I do this assignment, you*

know, and it really makes me want to do the assignment because I know it is not going to take as long and it's going to be fun…"

In the following section, I explore participant reactions in more detail, seeking to elucidate aspects of why participants liked LiquidText and how it impacted their experience for the better.

7.2.2 Why They Liked LiquidText

The details of participants' reactions to LiquidText are complex, and fall into several overlapping categories of response. So to explain why people liked the system, I discuss these results from three perspectives. First, I consider two areas of functionality that received especially positive feedback: comments and excerpts. Much of the feedback though, did not fit cleanly along functional lines, so second, I discuss two areas of user behavior where LiquidText seemed to perform especially well: content arrangement and navigation. Finally, I discuss several broad themes of advantages that recur throughout these data in an effort to understand what, fundamentally, was responsible for LiquidText's impact on experience.

Annotation

Several of the specific functions of LiquidText that participants saw as advantages align with the central processes of active reading; one example of this was annotation. Figure 7.1 Q5 shows participants strongly and significantly felt LiquidText was better at handling annotation than the control media (mean response = 4.08, where 1 is control and 5 is LiquidText). The interviews help explain why participants felt this way. P15, for example, pointed to improved speed as the benefit of LiquidText's highlighting and note-taking over paper. But more generally, of the five participants who discussed annotation as an advantage of LiquidText, three specifically cited the *connectedness* of comments, including their support for being interconnected to form groups, as well as their automatic

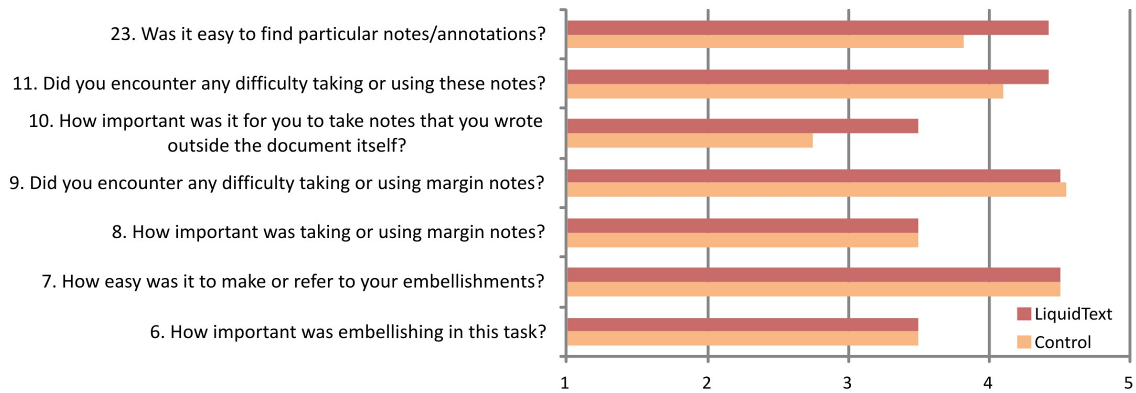


Figure 7.2. Mean responses from the general active reading questionnaire relating to annotation.

connections to their referents. P17, for example, focused especially on the challenge of post-hoc comment grouping using the control media:

> *"I feel like when I do tasks like this on paper I'll write my comments but it's really difficult to do something like this [as in LiquidText] where you can group things together because you can't really do that on paper, you can't group things around once you've done it."*

Closely related to grouping, two participants also discussed the freedom with which comments can be positioned as an advantage of LiquidText. On the other side of the annotation lifecycle, two participants also pointed to comment and highlight *retrieval*, noting the ease with which they could consult comments in response writing, as well as finding highlights in the text. The general questionnaires supported this, as participants rated ease of note/annotation retrieval 21%[7] higher in LiquidText versus the control (see Figure 7.2 Q23; difference is borderline significant with ANOVA at p=0.078).

The general and medium-specific questionnaires also supported the participant interviews more broadly. They confirm that it was not *overall* ease of use that made LiquidText's annotation support preferable. For both conditions, participants gave the same scores for the ease of making/using embellishments, and virtually the same scores

[7] Percentage differences are calculated after normalizing scores such that the lowest possible value is 0.

for the difficulty of taking margin notes (Figure 7.2 Q7 and Q9). Thus, the difference appears to have been in some of the sub-processes that occur under the broader heading of annotation—such as the above example of retrieval, or the examples of grouping. Notably, most of these and other advantages relating to annotation pertained particularly to comments rather than to highlighting. This too is reflected in the questionnaires, as participants on average felt that LiquidText's overall annotation support was more useful than its highlighting support in particular (means=4.48 and 3.5 respectively; difference was significant using ANOVA at p=0.03), thus leaving comments to account for the difference.

Excerpts

Complementary to annotation, many of the participants' positive reactions were to LiquidText's excerpting functions. Participants' comments about why they liked excerpts tended to fall into 4 categories: 1) retrieval, 2) arrangement, 3) excerpt connectedness, and 4) overall affect. Within the last category, participants discussed the *feel* of making and using excerpts. P21 said "And it's so simple and so—it's very pleasurable… [I] basically feel like I'm taking it for myself, you know. This is a good feeling." P17 went further,

> *"So I felt like when I was [doing the control condition] it was like, why did I sign up for this; I don't want to do Microsoft Office for 2 weeks. Whereas with this [the experimental condition] I was like, oh this is fun and I could create a giant note sheet with my own little excerpts."*

The notion of ownership and control hinted at by P15 and P17 is reflected elsewhere in feedback on the excerpts, such as with P13, who discussed how they allowed him to consciously elide the parts of the document he did not see as relevant, effectively making a simplified version of the text from the excerpts he considered interesting.

Participants also discussed to several other positive aspects of excerpts. With retrieval, for example, participants discussed both the sense of safety that excerpts cannot accidentally be lost, as well as the convenience of being able to easily browse one's excerpt collection. Another positive aspect was arrangement, as participants appreciated being able to organize excerpts together into groups, and generally appeared to find such grouping to be easy to perform. And while the interviews did not raise the issue of excerpts' impact on efficiency, in the LiquidText-specific questionnaires, participants felt excerpts had a large effect on helping them to be efficient (mean response = 4.3, where 5 is "more efficient" and 1 is "no difference"), suggesting another way excerpts contributed to participants' experiences.

But in spite of the positive responses, two of the tasks that I expected excerpts to support, note taking and outlining, were seen as only marginally better in LiquidText versus the control. As shown in Figure 7.1 Q17, the mean response to the question of which medium was better at supporting note taking and outlining was 3.67 (with margin of error = 0.95), where 5 was LiquidText and 1 was the control. But while this suggests LiquidText is only minimally, if at all, better at supporting outlining, one could argue that is not entirely bad. For example, P17 discussed that if LiquidText had outlining functionality per se, she would have used it out of habit, but she was glad it did not, as she thought the way she actually performed the task was better. As such, it appears that the contribution of excerpts to the experience of LiquidText was principally through other means, such as hiding irrelevant text, retrieval, and the like.

Navigation

While some of the positive aspects of LiquidText were associated with a specific function, such as annotation or excerpts, other positive aspects were broader, relating more to a *particular user behavior* and potentially involving several functions. One of these areas of user behavior was navigation.

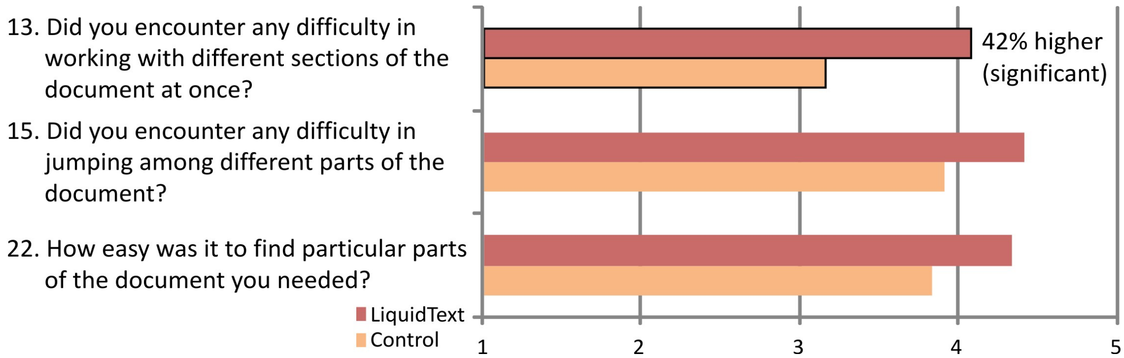


Figure 7.3. Mean responses from the general active reading questionnaire relating to navigation. Higher scores are better.

Navigation accounted for some of the most significant differences in participants' experiences of LiquidText versus the control condition. As seen in Figure 7.1, Q12, the comparative questionnaire showed participants found navigation to be significantly easier in LiquidText versus the control (mean response = 4.58 where 5 is LiquidText and 1 is control; margin of error = 0.42). A similar perception was found in the interviews, such as with P11 who said, "I can just skim through the entire article and come back to it and collapse and scroll whenever I want, so I found that more efficient than using Word." Two other participants expressed similar sentiments, describing how LiquidText saves one from having to shuffle or search through materials, in effect reducing the amount of navigation required when switching between different text objects. P9 felt the times when LiquidText would be better than the control were specifically those times when frequent jumping through a document was necessary.

The idea that LiquidText's navigation provides an especially improved experience when switching or jumping was partially reflected in the questionnaires as well. During the LiquidText condition, participants on average reported 18% less difficulty in jumping between different parts of one's document then during the control (Figure 7.3 Q15). The same was true for the ease of finding a particular part of the document, where LiquidText elicited an 18% higher average score as well (Figure 7.3 Q22). But while interesting, neither of these differences are significant with ANOVA at $\alpha = 0.05$. In contrast, the level of difficulty participants experienced when working with different sections of the

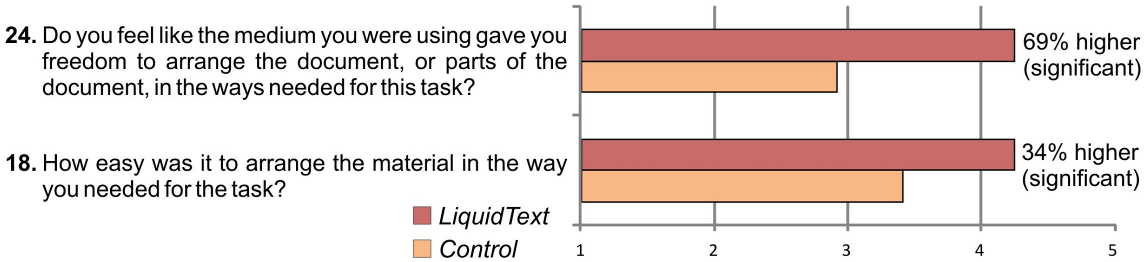


Figure 7.4. Mean responses from two questions on the general active reading questionnaire pertaining to arrangement. Higher scores are better.

document *at once* was significantly less (single factor ANOVA, $p = 0.02$), suggesting that parallelism is an important part of why participants preferred navigation with LiquidText (Figure 7.3 Q13). This was echoed in the interviews, such as with P11 who explicitly noted that comparison was easier with LiquidText than the control. P10 similarly felt that comparison—but for longer documents—is a strength of LiquidText.

Generally then, the improved experience of navigation with LiquidText appears partially to be based on efficiency, that the user can simply move from one place or tool to another quickly and easily. But an additional, perhaps more significant, component is in parallelism—allowing the user to navigate to and work with multiple pieces of text at once.

Content Arrangement

Besides navigation, another area of user behavior where LiquidText seemed to excel was content arrangement. And like navigation, content arrangement is a core part of numerous functions in LiquidText, including the positioning of the excerpts and comments, scrolling/collapsing of the document, and so on. And as discussed in Chapter 2, this is not idiosyncratic to LiquidText; the concept of arranging materials in space is an important part of active reading in general.

It is notable then that participants found LiquidText to be far better at supporting content arrangement than the control media. As Figure 7.1 Q14 shows, participants strongly preferred LiquidText for content arrangement, with a mean response of 4.5, where 5 is LiquidText and 1 is the control (margin of error = 0.57). This effect was supported by the general active reading questionnaires on two questions in particular (see Figure 7.4). On Q18, participants reported 34% higher scores for the ease with which they could arrange their materials during the LiquidText condition versus during the control (difference was significant; single factor ANOVA with $p = 0.049$). On Q24, participants reported an even greater difference when asked specifically about arranging their documents and pieces of their documents, with a 69% increase for the LiquidText condition versus the control (difference was significant; single factor ANOVA $p < 0.01$). Notably, the scores for these two questions (arranging materials generally, and documents/pieces specifically) were the same for LiquidText (4.25 out of 5). The reason Q24 shows a much bigger difference between conditions than Q18 is thus because of the control condition. In Q18, the control received a respectable average score of 3.4/5. But in Q24, the control condition average score was 2.92/5. Thus we see that while some arrangement tasks are manageable with the traditional PCs and paper, arranging full and partial documents represents an area of especially pronounced contrast between LiquidText and traditional media.

Beyond the quantitative measures, participants' qualitative feedback helped explain why LiquidText was better received for arrangement. In some cases, participant reactions were quite general, such as saying that LiquidText feels less cluttered than the control, and that it was, "nice to have a lot of work space here." Or P21's reference to the workspace,

> *"There is an intermediary state of the text, which provides an opportunity for structuring, rearranging and reflection (annotation) which was very useful in LiquidText, and completely missing in Word."*

Participants also repeatedly discussed the workspace as promoting awareness, as with P15, "…but I feel like I have more access to the information that I'm pulling out [in the

LiquidText condition] than I do in that sort of document, where I'm always hiding it somewhere."

And as noted in the previous section, participants also found arrangement valuable in supporting comparisons, and otherwise working with multiple sections of text together. Typically, participants appeared to use excerpts for this purpose, but in some cases collapse made a significant impact as well, as with P11, "And the collapsing, that's an amazing thing…That's the highlight of LiquidText I guess, where I can view two parts of the document together."

Besides comparison, another perspective from which to consider arrangement is organization. Here too, participants felt LiquidText was significantly better than the control, as shown in Figure 7.1, Q13. Specifically regarding the organization of thoughts to write the essays, participants preferred LiquidText strongly (mean response = 4.29, where 5 is LiquidText and 1 is the control; margin of error = 0.64). This was echoed in the interviews, as with P17 who noted, "Liquid Text is a lot easier to use intuitively…[and] a lot easier in terms of organizing and knowing where things are." Likewise, P11 felt that LiquidText enables a freedom of organization that is absent from Microsoft Word.

Throughout arrangement though, we see a repeated trend: that part of the way that arrangement conferred its advantages often came down to excerpts and annotations. As discussed above, participants noted how LiquidText allowed them to arrange their excerpts and comments so they could easily maintain awareness of relevant parts of text, as well as to organize their thoughts, and parts of the text itself, into groups. An important aspect, sometimes implicit, in many participant reactions was that comments and excerpts could be positioned arbitrarily. Again quoting P11, "…making notes was easier [with LiquidText]. I mean, I had something there like a workspace where I could position my notes and my critiques wherever I wanted."

Themes

In the above discussion, I explored the ways in which LiquidText's annotation and excerpt functions provided a preferable experience to their control condition counterparts. Likewise, I discussed some of the reasons that user behaviors like navigation and document arrangement were similarly perceived as better on LiquidText. But while some of the reasons for these differences were fairly idiosyncratic (for example, the fun of creating excerpts), there were certain common themes that emerged as well. The most significant of these was flexibility.

Perhaps most saliently, flexibility was found in the idea of spatial arrangement, as described above, and reflected likely the greatest difference between the experiences of the two conditions (see Q24 in Figure 7.4). Adding to the above discussions, participants provided thirteen comments directly, positively addressing the flexibility and structure of LiquidText. P11, for example, said that the workspace made her feel like she had more freedom while P15 put it more bluntly saying, "I can do whatever I want here [in LiquidText]." P17 elaborated by way of example,

> *"I think it's really nice and open ended…if I wanted things on top of each other or connected I could do that…If I wanted them in a circle I could have done that; I could have even done…when there is one thing in the center and then things connect out."*

P18 took a more complex view, affirming that LiquidText was both more structured and more flexible at the same time. Together the statistics and interview comments provide a consistent picture of a significantly freer, more flexible active reading experience using LiquidText than using the control.

A second theme that was repeated throughout the data was that of awareness—awareness of where one is in their documents, awareness of one's previous thoughts and notes, awareness of what has been completed and still needs to be done, and so forth. This appeared in several places noted above, such as with P13 who used excerpts to

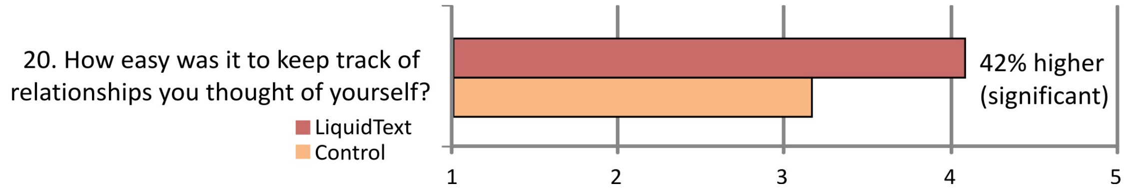


Figure 7.5. Mean response from question on the general active reading questionnaire pertaining to awareness. Higher score is better

reduce his awareness of irrelevant data, and using content arrangement to maintain an overall sense of access to one's data. But the interviews showed more about the role of awareness in the active reading experience. Akin to the experience of P13, P15 felt LiquidText allowed him to keep less data in his mind than in the control condition, again effectively allowing him to maintain mental awareness of less. P17, by contrast, felt that LiquidText's preview pane (Figure 5.1 B) helped her see clearly how long the document was, letting her maintain awareness of more. She additionally liked LiquidText's approach to notes, as it allowed her comments to remain more visible and accessible throughout her active reading task.

The questionnaires helped support this verbal feedback. Consistent with P17's comments, the comparative survey showed that participants found themselves considerably more oriented in LiquidText than the control (see Figure 7.1 Q11, mean response was 4.0, where 5 is LiquidText and 1 is control; margin of error = 0.72). Likewise, after the LiquidText condition, participants found it 42% easier to keep track of relationships they thought of during their task versus with the control (see Figure 7.5, difference is borderline significant with ANOVA at p=0.053).

Thus we see from the questionnaires how LiquidText helped readers maintain awareness of certain important information—such as their orientation. And likewise, from the interviews, we see that in some cases LiquidText allowed readers to reduce their awareness of things they preferred hidden. Cumulatively then, rather than merely increasing awareness, LiquidText supported a more selective cross-section of where the reader's awareness is maintained and where it is not.

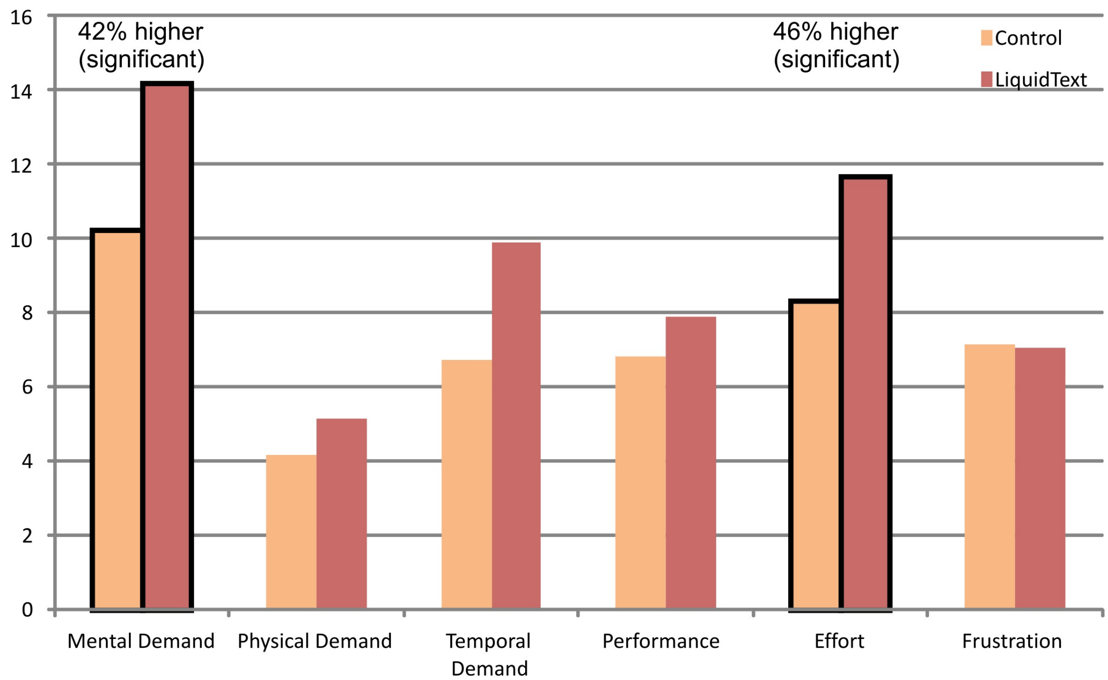


Figure 7.6. NASA TLX cognitive load test mean responses.

7.2.3 Why They Did Not Like LiquidText

As explained above, LiquidText's overall impact on the active reading experience was quite positive, particularly in terms of certain active reading tasks and processes, such as annotation and navigation. But there were other areas where LiquidText's influence on experience was not so clearly beneficial.

Perhaps the most surprising example of this is that, in spite of participants' strong overall preference for the system, LiquidText appeared to provide little, if any, overall reduction in the difficulty of the active reading task. For example, Q6, in Figure 7.1, shows that participants, on average, found active reading with LiquidText only slightly easier than with the control, and even this was well within the margin of error. Q9 asks which *medium* was easier to use, and is marginally more in LiquidText's favor than Q6, but still within the margin of error. This is strongly supported by the results of the NASA TLX cognitive load tests (Figure 7.6), which indicated that mental demand and perceived effort were significantly *higher* for LiquidText than the control (ANOVA; p=0.014 and *p*=0.028 respectively). Physical and temporal demand, as well as performance, were also nominally worse for LiquidText, but those differences were not significant.

These results are interesting: how can LiquidText be rated as providing an improved experience in many parts of the active reading process, and be preferred overall, yet be rated as only marginally easier, and as imposing a significantly greater demand on the user? The interviews help to explain this by showing places where the LiquidText experience was negative, as well as where the control experience was better.

Perhaps the most serious difficulty participants experienced while using LiquidText was in the gestures. While there were rare problems with several of the gestures (zooming, highlighting, ungrouping), most were very well received. The LiquidText-specific questionnaire even showed that participants rated the ease of remembering and performing the gestures at 4.6 and 4.2 respectively (where 5 is best). Indeed, the interviews echoed this, with most participant comments supporting the use of multitouch. Yet one problem was noted repeatedly: the selection gesture. In the LiquidText-specific questionnaire, 5 of the 12 participants noted selection as a difficult gesture—far more than any other. P13 even asserted that while using LiquidText, text selection was the hardest part of the entire active reading task, and P15 and P21 said specifically that they could select text better with a mouse.

Beyond selection, another problem may have been the hardware and spatial environment. As indicated by Figure 7.1 Q16, the legibility of text on LiquidText was considered only slightly better than the control, and this was well within the margin of error. Even P17, who generally strongly favored LiquidText, admitted that paper may have been easier to read than the display on the LiquidText tablet. Supporting this is the reality that text was necessarily rendered smaller on the tablet's 17" display than in the PC condition—and such size differences can affect readability [Bernard, Chaparro et al. 2003]. Size may have also played another role, as participants may not have had enough working space with LiquidText. In the general active reading questionnaire, participants

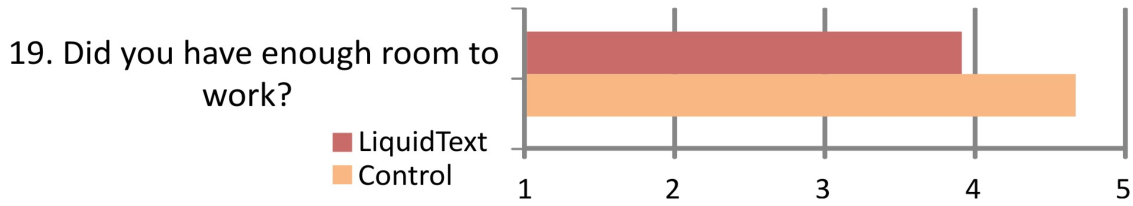


Figure 7.7. Mean response from question on the general active reading questionnaire pertaining to work room. Higher score is better

provided 20% lower responses to the question of whether they had enough room to work after using LiquidText versus the control (Figure 7.7). While the difference is not significant (ANOVA p=0.1), it may hint at another factor affecting the LiquidText experience.

Beyond the hardware, there were also suggestions of certain areas where the functionality of LiquidText itself potentially fell short. One of these, significantly, was that several participants felt they needed more expressiveness. For example, P11 wanted to be able to write notes in her own handwriting—a function clearly available in the paper environment of the control condition. P15 and P21 felt LiquidText needed ways to show higher level structures, such as complex relationships among groups of excerpts—though it is uncertain whether the absence of such functions could have made LiquidText more difficult than the control, as the control did not necessarily afford these things either. By contrast, at least four participants felt LiquidText needed a simple text-search function to help them find particular parts of their documents. This certainly was available in the PC environment of the control condition, and so could have relatively diminished the LiquidText experience.

More surprisingly, in contrast to the users who felt LiquidText needed to support more expressiveness, a small number of participants felt it needed more structure. P13, for example, felt documents should not move laterally when being dragged, and that the space between excerpts when manually positioned was wasteful. P15 likewise was concerned with the large range of color intensities in the highlighting control, as he was afraid he would not be able to select the same shade consistently. However, these were infrequent problems, and seem unlikely have played a significant role in the overall experience.

However, a problem that could have had a broader influence on experience was training. The experimental setup included approximately three hours of practice with LiquidText, but it seems unlikely that this could provide the fluency of the years participants spent using media like that of the control. Nonetheless, from an experiential standpoint, there was little data to support this supposition. P10 noted, "When I read paper, it's easier for me because that's what I usually use…" but beyond that, participants seemed to have felt comfortable with the technology. Indeed, when asked if they had enough time practicing using LiquidText, participants gave an average response of 4.5 (where 5 is best). Thus, while lack of experience may have impacted participants' use of LiquidText, there is little clear evidence that it played a significant role in their experiences.

Another possible problem could have been LiquidText's use of touch. Generally, the use of multitouch is still fairly new, and its use in knowledge work tasks still unusual. Yet, participant reactions to multitouch were generally quite positive. While the questionnaires did not address the issue of touch generally, three of the four participants who discussed it in the interviews felt that the types of tasks included in the study were suited to touch. P18, for example, felt it was, "just easier to just go like this—use your finger for everything." P13 agreed, saying these types of tasks are well suited to touch, except for text selection. The dissenter, P15, even felt LiquidText's gestures "were good," but was unsure whether touch would be better than a mouse (again, apart from text selection). So apart from selection, LiquidText's non-traditional application of multitouch does not appear to have been a problem.

In contrast to the previous section, there do not appear to be broad themes overarching the negative aspects of the LiquidText experience; rather, feedback from the questionnaires and interviews suggests the problems were more ad hoc. The hardware, being the main example, seems to have been inadequate for the task, leading again to the age-old computer mediated active reading problems of a lack of screen space [O'Hara and Sellen 1997], and poor legibility. And while participants generally were comfortable with the gestures, text selection seems to have been substantially detrimental to the user experience. Together, these two problems, along with a medley of small feature requests, is likely what led LiquidText to present a more challenging experience than the control.

7.2.4 Answering the Research Question

The third research question asks about impact of a computer mediated, highly flexible, high degree-of-freedom text representation on the subjective experience of active reading. To answer this question, we can consult the above discussion and look at participants' positive and negative reactions to the different aspects of LiquidText. It is thus tempting to conclude that the answer to the research question is complex and mixed, with positives and negatives—a statement which is certainly true of LiquidText itself. But I propose that the impact of the underlying flexible document representation was more consistent.

If we consider the functions of LiquidText which were most lauded—which seemed to have had the most positive impact—they included the annotation, excerpting, navigation, and several content arrangement functions. But the *attributes of* these functions which participants gravitated towards are telling. Recall the above theme of flexibility: participants pointed repeatedly to the workspace, and its support for complete and arbitrary control over object positions and grouping, and over arrangement in order to see multiple pieces of content in parallel. Similarly for the theme of awareness, participants appreciated being aware of more, but also of less, when they had more control over what they had to be aware of. So while participants generally liked the use of multitouch and the physical-like behavior of LiquidText, the functions and attributes that had the strongest positive impact were those that emerged directly from the philosophy of providing a high degree-of-freedom, flexible representation.

By contrast, the problems with LiquidText were not trivial, but neither were they grounded in the design philosophy behind the system. The limitations of the hardware, for example, could be solved without changing LiquidText's design philosophy at all. Likewise, the challenges of the text selection gesture are more complex but even those could likely be solved with better touch sensing hardware [Benko, Wilson et al. 2006].

Cumulatively then, I argue that the flexible document representation on which LiquidText is based leads to an improved active reading experience. While we know from the TLX surveys that this does not generally mean an overall *easier* experience, the comparative questionnaires suggest that it does mean a more enjoyable, and indeed preferable, experience, compared to the typical media used in the control condition.

7.3 Active Reading Process

The purpose of this section is to explore the impact of LiquidText's flexible, high degree-of-freedom representation on the *process* of active reading. The first aspect of this is to directly consider how readers' processes differed when performing the LiquidText condition versus the control. For this, I explore apparent process differences which I identified through analysis of the videos, participants' workspaces, and their interviews.

The second aspect of this section is to consider the extent to which LiquidText's functionality was appropriated. Here, I discuss the major areas of the system and the role that features played in the reading process.

7.3.1 How Behavior Differed Across Conditions

Understanding LiquidText's impact on user behavior is a multifaceted problem. Because the functionality offered by LiquidText is so different from much of what was available in either of the control condition configurations, there are necessarily differences arising from the mere fact that certain affordances of paper or PCs were not present in LiquidText, and vice versa. And while I did identify some such differences, there were, remarkably, larger scale impacts of LiquidText where readers *could* have behaved similarly to the control, but chose not to.

I sought to understand LiquidText's impact on the reading process in several ways, but some of the most direct data came from the interviews. There, participants were asked after each session to explain how they performed the reading task—the steps they took, the materials they used, and so on. In these sequential descriptions, participants generally described taking actions such as reading the article, highlighting, and iteratively answering each part of the prompt. But the LiquidText condition tended to include another action—creating a separate, intermediate document that was derived from the article text, and to which the participant referred while writing her response. This is essentially a notes sheet, but it has the particular property that participants tended to refer exclusively or almost exclusively to *it* when creating their responses.

Less abstractly, P10, P11, P13, P17, and P18 discussed their control condition strategy in the interviews. Of these, all (with the partial exception of P17) described how they would read the article, annotate it in various ways, and then write their essay, consulting the annotated article as they went. Of course there were individual variations within this pattern, such as participants who performed all of their annotation as they read through the article (e.g., P17), versus others who made multiple passes to apply different types of annotation (e.g., P18). But this is in contrast to the processes used in the LiquidText condition, which were described by six participants in the interviews. Of these, P13, P15, P17, and P18 clearly discussed using at least one of their passes through the article to pull excerpts or comments into the workspace, and then principally consult that separate collection of materials when writing their responses. There were also individual variations in this case, such as P18 who created excerpts as he read the article, in contrast to P15 who used one pass to read and highlight and another to read his highlights and make excerpts. Ultimately though, the participants generally included this

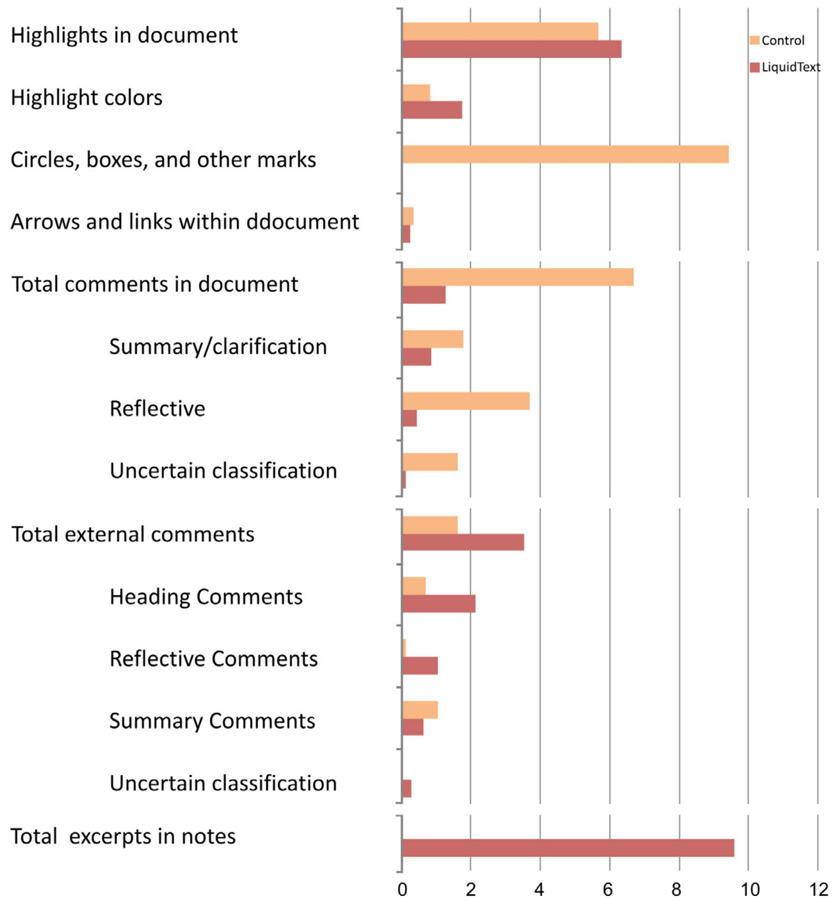


Figure 7.8. Average number of annotations of different types made in the LiquidText versus control conditions. Note: "in document" refers to annotations in the original article, whereas "external" refers to annotations on a separate surface.

intermediate document, in the form of materials brought to the workspace, as a core part of their reading process.

And this use of an intermediate document with LiquidText naturally tended to result in different types of notes and annotations in the control condition versus the experimental. For example, participants created an average of 6.7 comments within the article itself in the control condition, versus 1.3 for LiquidText. By contrast, comments written on a *separate surface* (e.g., a note sheet, or the LiquidText workspace) had the opposite relationship, with participants creating an average of 3.5 with LiquidText and only 1.6 with the control[8] (see Figure 7.8). Note that, with an extreme outlier removed from the latter statistic, both of these differences are significant with a single factor ANOVA, with $p < 0.05$ in both cases. Likewise, the same pattern is seen for excerpts.

Even though both LiquidText and the PC configuration of the control condition supported excerpting in some manner, participants using LiquidText created an average of 9.6 excerpts, whereas in the control condition they created zero[8]. Thus we see that in the control condition, participants tended to annotate the article itself, whereas using LiquidText, they preferred to create a separate arrangement of notes and excerpts external to the article.

Beyond textual notes like comments and excerpts, non-textual marks reveal a similar pattern—albeit for a very different reason. Hand-drawn annotations including the use of underlines, circling or drawing boxes around text, etc. is only possible in the paper + PC configuration of the control, and participants made an average of 9.4 such marks in the control condition overall. LiquidText, of course, had zero. But while LiquidText does not support hand-drawn marks, participants appeared to use several other behaviors to achieve the same effect. The first was the use of excerpts. Although excerpts are rich, textual notes, with very different visibility and organizational properties from hand-drawn marks, they were still sometimes used comparably, to raise the salience of a piece of content, as described by P13 and P15 in the interviews. P15 explains, "I liked highlighting a lot and then…basically going through and kind of highlighting my highlights by dragging them over here [to the LiquidText workspace]. So I've kind of gone through two levels of importance, I guess." Here P15 goes beyond just using excerpts to raise salience, but uses it in conjunction with highlighting designate different levels of importance. Similar behavior was also seen in the control condition by participants who used different combinations of highlighting, circling, underlining and outright note taking to designate these levels.

[8] One participant did create a note sheet during the control condition that she used for summarizing the beginning of the article. However, she deleted it before we could save it, so it is possible she created excerpts or comments that we missed in reviewing the video recordings.

Notably though, it was not clear whether participants used LiquidText's highlighting functions to supplement other, missing, forms of marking. This is visible in Figure 7.8, as average numbers of highlights created with LiquidText were only marginally greater those created in the control condition (a difference of about 11%). However, in LiquidText participants did use additional colors of highlighting (1.75 versus 0.83 per participant on average), which may have acted as a supplement for the added richness of hand-drawn marks. Unfortunately the interviews shed little light on this, as participants explained their uses of highlighting but did not suggest that it related to the absence of any other LiquidText functionality. So ultimately, a complete understanding of participants' strategies for coping with the lack of hand-drawn notes in LiquidText is beyond the scope of these data, but may be possible in future studies of the system.

Meaning

While the above discussion helps explain which types of annotations and notes were used, in some cases we can also consider the purposes that these notes served. An

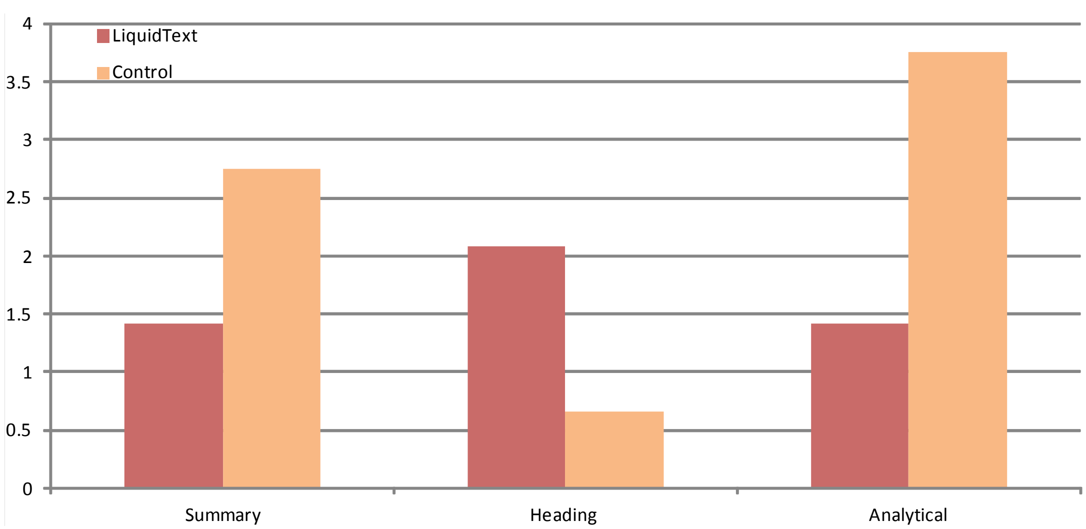


Figure 7.9. Average number of comments that were either analytical, provided a summary, or acted as a heading in both conditions.

example of this is seen in the different types of comments that participants created. To explore this, I classified the comments made in both of the control condition configurations, as well as in LiquidText, into three categories. The first is *summary*, which refers to comments that recapitulate or clarify some existing material, generally the article. The second is *heading*, which refers to comments that act as a title or tag for describing some other group of notes. The final category is *analytical*, which includes comments that reflect upon, critique, or opine about some existing material, again typically in the source article. As shown in Figure 7.9, the distribution of these three comment types varied widely between the two conditions—notably though, these differences were not statistically significant at the $\alpha = 0.05$ level; the closest was analytical comments, with $p = 0.11$. With one extreme outlier removed, heading comments were also borderline significant, with $p = 0.08$. But while we cannot be certain of their generalizability, these differences do align with several larger trends in how people used LiquidText at least in the case of this study.

One of these trends related to the use of heading comments. Because participants created large numbers of excerpts during the LiquidText condition, one aspect of their reading process was spatially arranging these excerpts into groups which, for example, in some cases addressed different aspects of the source text. Participants sometimes distinguished these groups using comments as headings, which was possible through LiquidText's functions for creating comments that are not linked to a referent. Because in the control condition, participants generally did not create a separate, intermediate document, their ability to organize their comments or excerpts was minimal, and so it is unsurprising that heading comments were rare in the control.

But perhaps a more striking feature of Figure 7.9 is that both summary and analytical comments dropped considerably in the LiquidText condition. But by itself, this is misleading. Summary comments did drop, to be sure, but that does not mean that less summarization was being performed, since various other functions were used for summarization as well, such as excerpts. P18 described this, saying he would pull out excerpts corresponding to the main ideas of the article; and P13 likewise described using excerpts as a way of capturing a thought within the article but without the irrelevant parts of the text. This was also true of other forms of annotation, including highlights, which were used like excerpts to capture parts of the article that could serve as a summary. But while several devices were used for summarization, analytical comments appear to stand alone—that is they do not appear to be readily interchangeable with other devices. As such, an alternate way to view the reading process in the two conditions is to separate analytical comments from all other annotations/notes and look at them in parallel, as shown in Figure 7.10. This figure shows that, excluding analytical comments, total notes

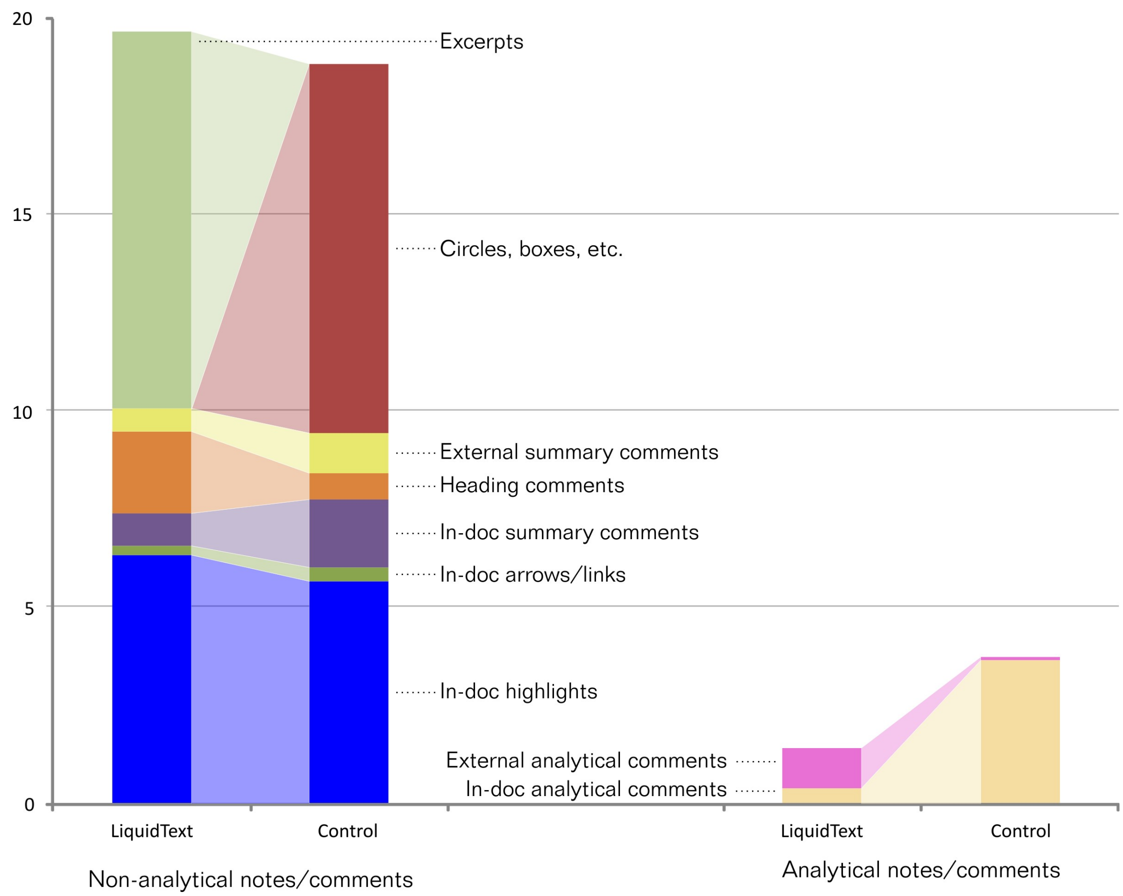


Figure 7.10. Average number of analytical and non-analytical notes and comments made in the LiquidText condition versus the control condition.

and annotations were quite comparable between conditions. Indeed, LiquidText seemed to make up with excerpts what it lost in freeform markings compared to the control. Likewise, participants tended to make similar numbers of highlights in both conditions as well. Non-analytical comments likewise were used with similar frequency between conditions, with the increased use of heading comments with LiquidText making up for the decreased use of summary comments.

On the whole then, the total use of marks, annotations and notes was comparable between conditions; the only major outlier was analytical comments, which appeared far less frequently in LiquidText. But the reason for this is not entirely clear, as even P15 described uncertainty about the cause of his behavior,

> *"But in a paper one and in a PDF version, I tend to just reference the text and make my critique right there, my comment. In [LiquidText], what I'm doing is I'm pulling out the comments right there and I'm doing the critique in my head. Wow, I never thought about that."*

However, one possible explanation may be related to the nature of the intermediate document that participants tended to create in LiquidText. That is, as discussed above, participants tended to use excerpts and comments to create a new document in the workspace, derived from the original article, which they then principally consulted as they wrote their response essays. Notably, these intermediate documents were summary-like in nature. We see this first in that they were predominantly composed of excerpts, which tend to afford a summarization function, and we also see this in the interviews. P13, for example, was explicit about this, describing himself as reading the document while constructing a summary of it. P15 gave a similar explanation, saying he would highlight and create excerpts for key points in the article. P7 too, described using the workspace to create a separate summary of each of the two positions described in the article. Thus, as exemplified by the quotes, and demonstrated more broadly through the statistics, we see that the LiquidText reading process generally entailed the creation of an intermediate, condensed, summary document, rather than written, analytical reflection on the article. I suggest that the creation of this intermediate summary document may have dissuaded participants from performing additional analysis—either because of lack of time or lack of motivation to perform additional reading process steps.

While conclusively deciding between these two options is beyond the scope of my data, there are hints that time was not the problem. Of the 11 participants for whom we have video data, 6 appear to have had at least 8 minutes remaining when they finished in the LiquidText condition—compared to 7 participants in the control, suggesting that in both cases, participants were generally not lacking for time. Likewise, TLX scores for

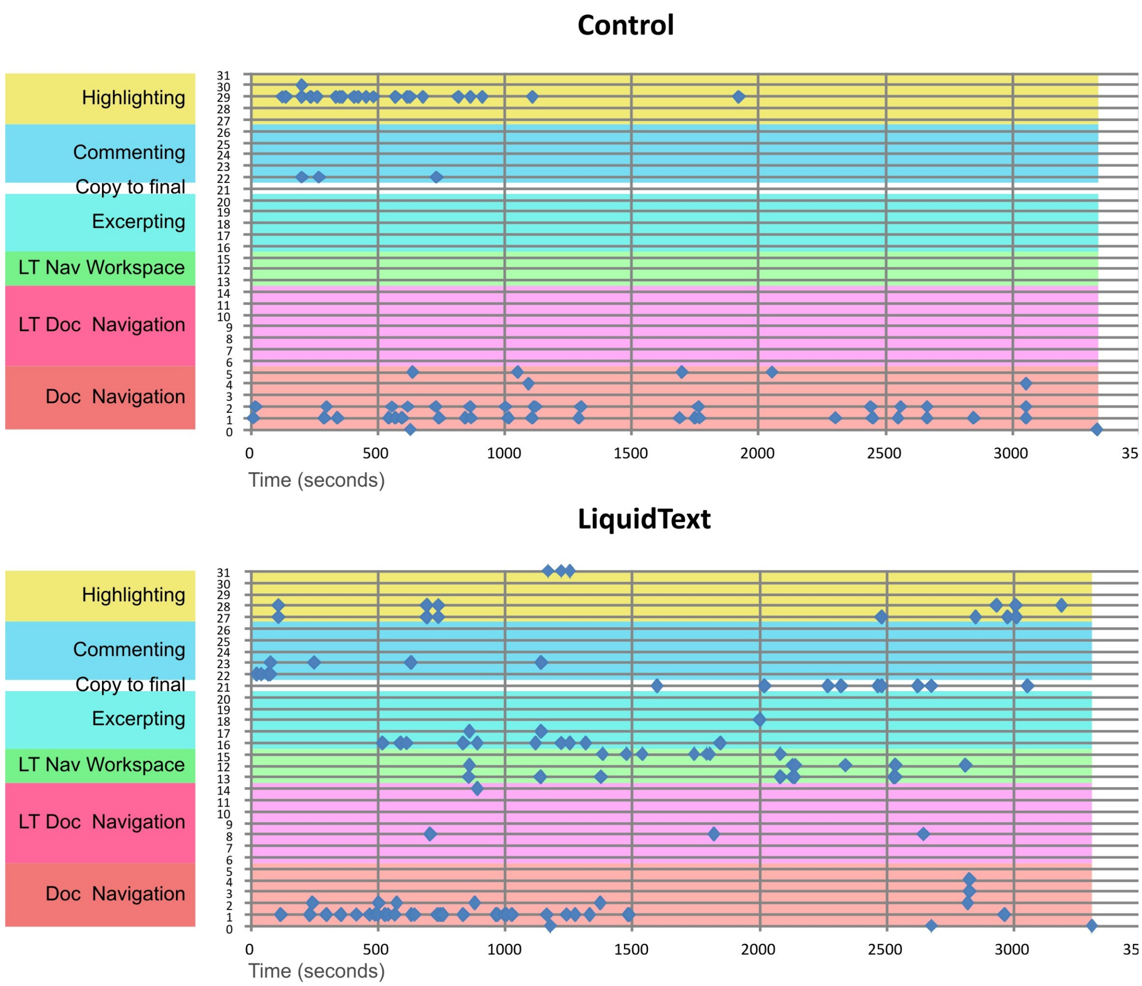


Figure 7.11. P2's timelines for the LiquidText and control conditions. Each numbered horizontal line corresponds to a different type of event; event types are organized into groups and color coded. See Appendix D for the specific functions associated with each number.

temporal demand were not extremely high when using LiquidText (mean of 9.9/16), and in any case were not significantly different from those of the control condition (mean of 6.8/16). If however, it was not for lack of time, it may have been for lack of motivation, as participants may not have felt that an extra, written, analytical step was required to perform a deep analysis of the article. As shown in Q8 in Figure 7.1, with whatever process changes participants engaged in during the LiquidText condition, they felt that on average, their thinking processes were better than with the control.

Behaviors over time

If it is true that while using LiquidText, analysis of the document was taking place in participants' heads as they consulted their intermediate text, then one would expect participants would likely apportion their time differently vis-à-vis the control. And we see examples of this in the distribution of navigation events during the reading process. Figure 7.11 shows as an example the activity timeline for participant P2 for both the LiquidText and control conditions. The timeline shows the occurrence of 31 different types of events over the course of the reading process, with event types grouped into color-coded categories. The timelines for all participants, as well as the specific events that the numbered rows correspond to, can be found in Appendix E.

One of the most striking differences between the control and LiquidText conditions for the above example of P2 is in Doc Navigation (short for document navigation). In the control condition, navigation events are spread out fairly uniformly over the duration of the reading task. But in the LiquidText condition, document navigation events are tightly clustered together into the first half of the timeline. Albeit somewhat less tightly, the comment and excerpt creation events are also clustered mostly in the first half of the timeline for LiquidText. By contrast, LT Nav Workspace (navigation of the LiquidText workspace) events were most frequent in the second half of the timeline. Thus, we can see that in the control, P2 tended to work from the original

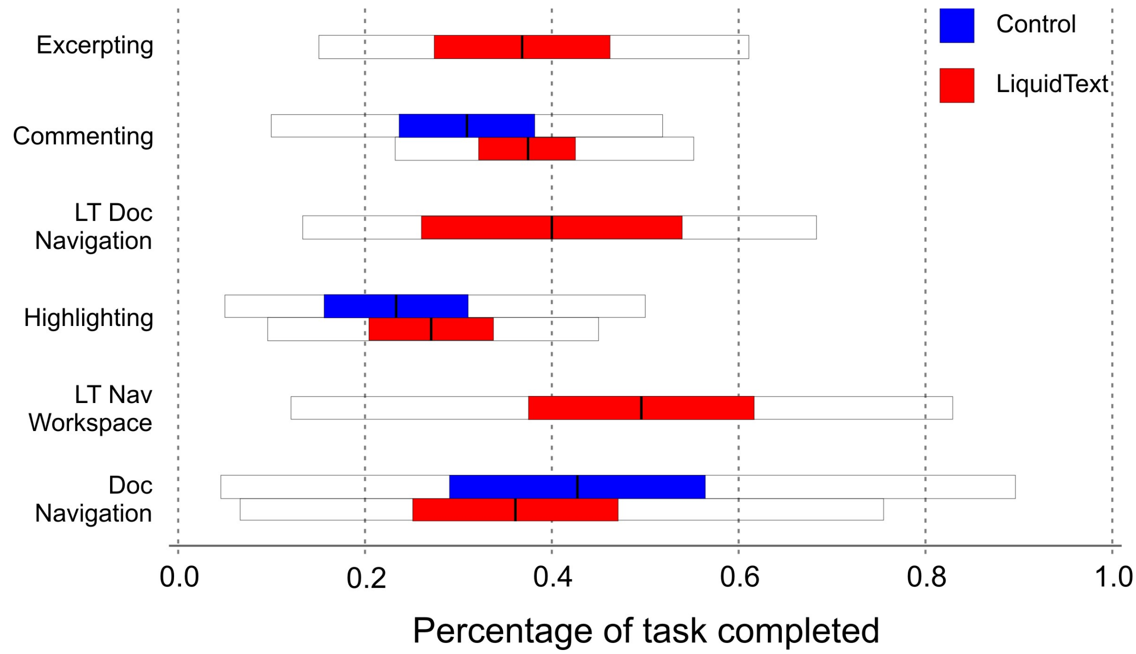


Figure 7.12. Overview of times at which different types of operations were performed over the course of the active reading task. For each category, the black vertical bar is the average of all participant means, the colored rectangle shows the average standard deviation among participants, and the uncolored rectangle shows the average max and min times among participants.

document. Even though highlighting and annotation were generally only performed during the first half of the task, P2 kept navigating and working with the original document throughout. But in the LiquidText condition, the behavior changes as we would expect. During the first half, the reader is navigating and working with the original document, as well as creating excerpts and comments. But during the second half of the task, the reader abruptly stops working with the original document, and instead navigates the LiquidText workspace, where the excerpt-based intermediate document is located.

And while these event distributions are especially clear with P2, several of them are present to varying degrees among the other participants as well. We see this in Figure 7.12, which provides an overview of when participants performed different categories of behavior. Figure 7.12 was constructed by calculating the minimum, maximum, average, and standard deviation of the points in time at which each participant began tasks in each category. These values (max, min, standard deviation and average time for each category) were then normalized relative to the total time the participant spent in the reading task, and then averaged over all participants. The results are shown in the figure. Note that participants who performed less than two actions in a given category were excluded from that category's averages.

Thus, Figure 7.12 echoes the patterns we saw with P2 in several ways. First, as with P2, document navigation tasks were generally performed over a shorter duration of time, and earlier in the overall reading task, in the LiquidText condition versus the control (the average max times for document navigation tasks were significantly different between the two conditions—single factor ANOVA, *p*=0.03). Similarly, highlighting, excerpting, and commenting in LiquidText, as well as LiquidText-specific document navigation operations (LT Doc Navigation) were all, on average, performed before workspace navigation. And as with P2, highlighting and commenting were also typically performed relatively early in the control condition as well. Thus the core difference we see between the conditions, as with P2, is that navigation in the control condition generally consisted of *document navigation* throughout the task, while with LiquidText, navigation changed from document navigation early on, to workspace navigation in the

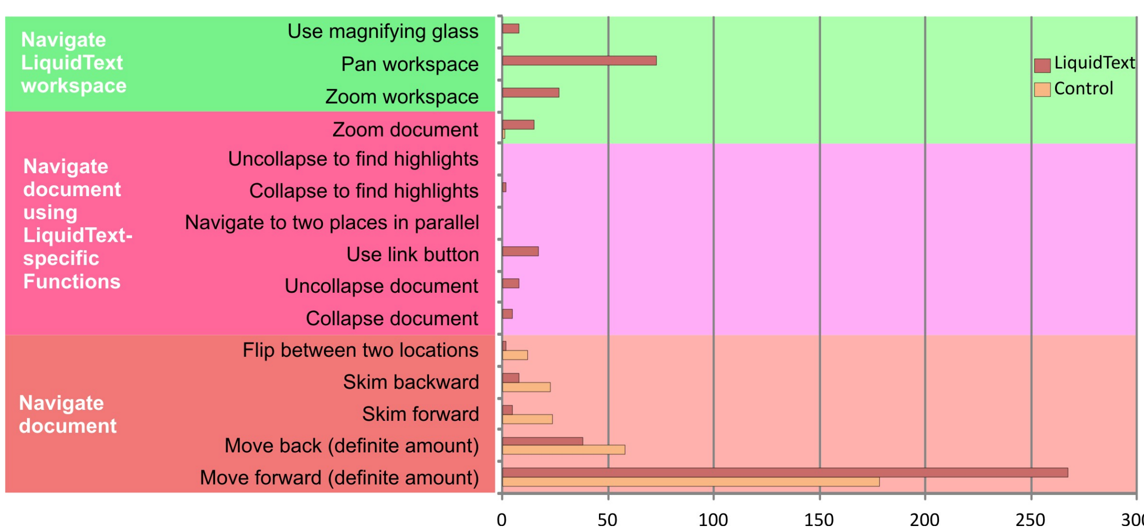


Figure 7.13. Aggregate histogram for different event types for both conditions. Events are color coded by category.

later part of the task. Therefore, while the differences between the control and LiquidText conditions are especially pronounced for P2, we see they occur broadly among the participants. Together they exemplify another manifestation, and help us understand the implications, of the reliance on an intermediate document while using LiquidText.

But while there is a change in the time distribution of navigation events with LiquidText, there is also a change in the types of navigation events that are used. First, as Figure 7.13 shows, participants actually performed more navigate-forward operations in LiquidText than the control—apparently a result of their making many small forward movements. The reason for so navigating is not clear, but may have been related to the shorter distance between participants and the touch screen than the traditional PC, and so more movements would be needed to keep the document in a comfortable viewing position. But given this propensity to make many small movements, it might initially seem surprising that there were fewer of all other document navigation events on LiquidText than the control. But in the context of an intermediate document, it becomes understandable. Forward movement is of course necessary simply to read the document, but backward navigation, skimming, and flipping are all part of non-linear navigation, and as such are involved with the more analytical, critical components of reading. And as we have seen, in LiquidText these parts of the reading process tend to be performed using the workspace more than the document. So in essence, it appears that in LiquidText, non-linear navigation events occurred in the "Navigate LiquidText Workspace" category rather than the "Navigate Document" category.

Why Behavior Changed

From the above sections, we see a variety of ways in which user behavior differs between the LiquidText and control conditions. We also see that generally, these differences appear to emanate from the use of an intermediate document. But the question of *why* users chose to create and consult a sort of summary of their text in LiquidText is less clear—especially since creating such a summary document would be possible in both configurations of the control condition as well.

The participant interviews did not directly address the question of why they chose to create intermediate documents, but nonetheless suggest some possible explanations. One is that LiquidText's user interface makes summarization easier than other note/annotation practices, and so people choose it by default. There are hints of this in participants' comments about LiquidText's prime summarization function: excerpting. P21 for example, said of creating excerpts, "And it's so simple and so—it's very pleasurable…I basically feel like I'm taking it for myself, you know. This is a good feeling." P13 went further, "Whereas here, because the text was already in editable format of pretty trivially being able to make…my [excerpts] and make that my summary without retyping stuff, which led to me doing it more, because I could." As discussed above in this chapter, other participants also had positive reactions to excerpts, such as P18 describing them as easier than writing out excerpts, or P7 saying the excerpt gesture is easier than copy/paste on the desktop. But the quotes by P21 and P13 are especially interesting because they go further, suggesting that excerpting was 1) especially enjoyable and 2) so comparatively easy as to promote, rather than just allow, summarization. Even though participants also noted the ease of commenting (such as with P11 and P7), the response to excerpts was simply much stronger.

But if it is true that people tended to use LiquidText based on which features were most easily accessible, this would not be surprising. Prior research in the use of behavioral economics in persuasion found that people regularly let small differences in convenience influence significant daily decisions [Lee, Kiesler et al. 2011]. Differences in affect can have a similar influence, as with the so called "Piano stairs" experiment. There, passengers exiting a subway terminal tended to use an escalator rather than a staircase—until the staircase was modified to look like a piano and play musical notes when stepped upon[9]. After the change, people took the stairs more often. Of course, it is unclear whether these effects hold true generally in active reading strategy, and still less certain if this explains why people tended to prefer excerpts over comments. Nonetheless, it raises an important question for future research, and potentially opens an avenue for guiding the reading process itself. That is, many books have been published that aim to teach people how to effectively read (such as [Adler and VanDoren 1972]); so it may be possible to encode such best practices into the UI of active reading tools themselves in a way that invites, but does not force, the user to comply with them.

Alternately though, there are several other factors that could have influenced participants' strategies for the reading tasks. One is the training they received. As I taught participants to use LiquidText, the training scenarios we stepped through tended to include the creation of several excerpts and only one or two comments. The reason for this was purely practical: I taught participants how to use the workspace after introducing the excerpts but before comments, and so they needed to make more excerpts as part of the tutorial on using the workspace. It is possible, though, that as a result, participants were primed to think of the workspace as a summarization rather than reflection area. A second, minor factor could just be the presence of the workspace in the LiquidText interface. As a large, empty region, it is conceivable that participants tended to feel that they should put something in it. Because excerpts are relatively easy to create, that could have seemed a more natural choice since it could populate the large, empty workspace more efficiently. Ultimately though, both of these possible factors are speculative, but they could point toward future research directions to better understand LiquidText's effect on the active reading process.

---

[9] A movie is available at http://thefuntheory.com/piano-staircase

7.3.2 Appropriation of LiquidText Functionality

Besides understanding LiquidText's impact on the reading process, the second component of my fourth research question is how participants appropriated its functionality. Of course, in addressing the first aspect of RQ4, I discussed several factors concerning how comments and excerpts were used and hinted at the applications of other functions as well. This section, however, goes into greater depth exploring which

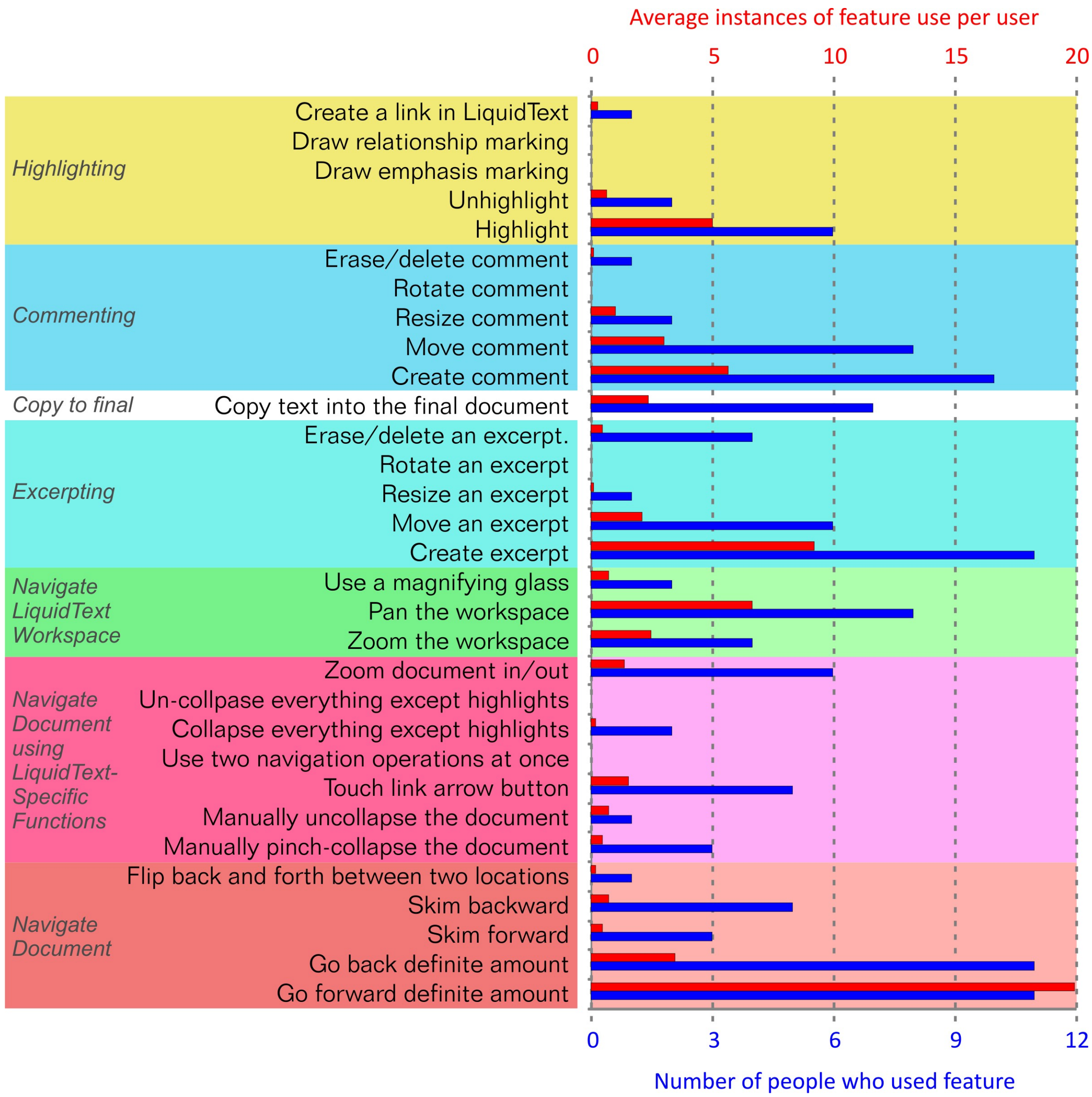


Figure 7.14. Chart showing A) the average number of times each user used each LiquidText feature (in red) and B) the number of users who used each LiquidText feature (in blue). Note that one user's video data was not logged, so the maximum numbers of users shown here is 11.

functions did and did not tend to be employed in the active reading process, as well as what purposes they served.

Some of the most readily appropriated functions in LiquidText were those relating to the workspace. We see this in Figure 7.14, which shows both how many times each participant (on average) used each function, as well as the number of unique participants who used it. Here we see that, besides scrolling forward, which was required to read the text at all, the most commonly used functions were those related to populating and navigating the workspace. This includes creating excerpts and comments, as well as workspace panning. Recall that even highlighting, which was employed by 6 participants, was often used to identify important text, a subset of which was then sometimes excerpted. Correspondingly, use of highlights and creation of excerpts/comments in the workspace were significantly correlated (Pearson $r = 0.61$). More subtly, though, the freeform nature of the workspace, allowing materials to be positioned arbitrarily, appears to have been leveraged as well. Participants described a rich variety of workspace arrangement strategies, such as P15, who organized his materials into about six groups corresponding to various concepts he found in the document, with several ungrouped comments/excerpts floating individually. P17 organized excerpts more chronologically. Others were more arbitrary, as with P18 and P21 who positioned excerpts randomly, though P21 placed associated comments in their proximity. Alternately, P7 described organizing his materials into two columns, corresponding to the two viewpoints he found in the article, and P13 similarly described a two-column, spreadsheet-like approach. And while there is insufficient data to identify conclusive categories of organizational schemes, there are already hints of a top-down approach (as with P17, whose organizational strategy is largely independent of content) versus a bottom-up approach (such as with P15, whose groups emerged directly from the article itself).

Considering comments in more detail, there was a general interest in *not* keeping comments attached to the article. This can be seen in Figure 7.14; the move-comment function was not used frequently, but was employed by most participants. The interviews reflected this as well. P15 explained that he was hesitant to leave comments attached because then he couldn't see the comment when he scrolled away. P21 went further, not even creating comments for the document itself, but explaining that it just felt more correct to comment on excerpts instead. And as described above, comments were generally used for summarization, reflection, or heading, with heading being the most prominent.

Also as described above, excerpts served several purposes, but were generally used for some form of summarization. Considering them in more detail though, we see that excerpts were sometimes used in place of, or as an addition to, highlights to view select pieces of text at higher saliency than the rest of the document, as with P13 and P15. P13 specifically pointed to excerpts as a means of supporting comparisons. P18 described a similar idea, explaining that they allowed him to see different areas of the document without having to scroll back and forth. This is notable because LiquidText includes separate functions dedicated to supporting comparison (collapsing), and particularly for highlighted text (collapsing to show all highlights). These functions, however, were each used by only three or two participants respectively during the reading task.

Overall then, we see that functions integral to the creation and review of the intermediate document were broadly appropriated into users' active reading processes. But many functions outside of this scope, including several of LiquidText's more characteristic, atypical operations, were used less often. We see this clearly among the LiquidText-specific document navigation functions. Collapsing, for example, was used by only three participants—one more than the number who used the highlight aggregation function. Others, such as using multiple navigation operations in parallel (such as holding the document in place while scrolling, or pressing two arrow buttons at once) were completely unused. The strongest exception to this, though, was the arrow

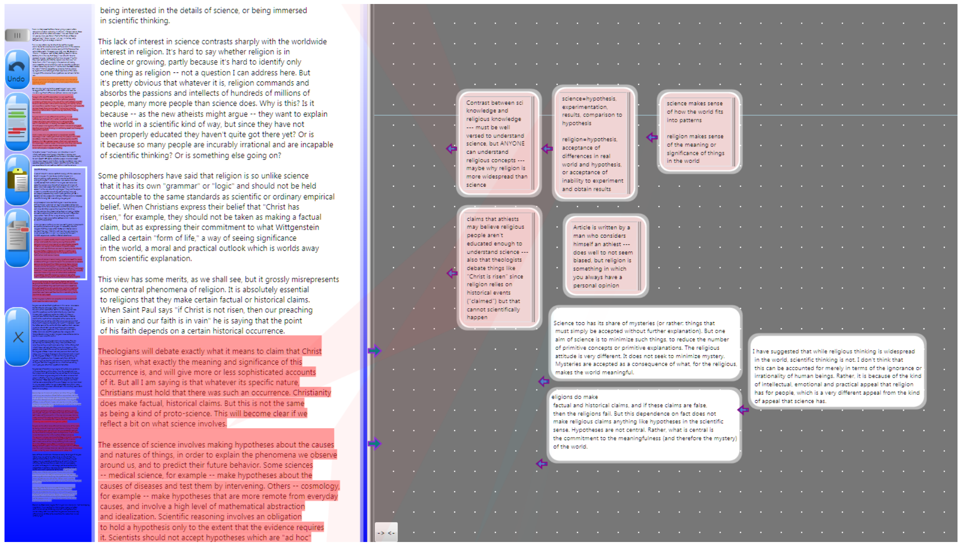


Figure 7.15. The LiquidText workspace of P17. Note that all of her comments and excerpts comfortably, and readably, fit on the screen at once.

buttons. Used one at a time, these were employed by five participants, but even then, only 1.5 times on average per study participant.

The reason for the low usage of the advanced navigational features is uncertain, but does not appear to be rooted in ignorance. After the training, for example, all participants were able to use the collapse and highlight retrieval functions, as well as the arrow buttons. Rather, the available data hint that these features were largely unneeded for this task. In the interviews, P13, P15 and P18 described using excerpts as a means of comparing passages, rather than collapse. P18 also described not needing the arrow buttons because he remembered where he put materials, or could just use the preview pane to identify the highlights in relevant areas of the document. Likewise, even without using these functions, participants still had very favorable reactions to the document navigation functions in the questionnaires, as discussed above. For example, the ease of seeing multiple parts of the document at once was rated 4.1 out of 5. Likewise, the ease of finding particular parts of the document was rated as 4.3 out of 5. Thus, participants appear to have been satisfied with navigation using LiquidText without the need of the features they did not use. But one possible reason these features may have been unneeded is that participants generally performed their analyses using the intermediate document they created in the workspace, and so functions to explore and visualize the original document would have been less necessary.

However, if it is true that participants found little need for the document navigation functions because they relied more on their intermediate document, then one would expect greater incorporation of *workspace* navigation functions into their reading processes. And indeed, as shown in Figure 7.14, the workspace navigation functions generally did show greater appropriation—e.g., the LiquidText-specific document navigation functions averaged 2.4 users apiece, whereas the workspace navigation functions averaged 4.7. The most frequently used of these functions, as noted above, was simple workspace panning, which was used by 8 of the participants (an average of 6.6 times per user). The other functions, however, were not incorporated as readily. Workspace zooming was employed by 4 participants with an average of 2.5 uses per participant. Moreover, magnifying glasses were used only by two participants. But the reason for the infrequent use of these two functions, though, is not clear. However, as with document navigation, it does not appear that this lack of use was associated with difficulties navigating the workspace. For example, users on average rated the ease of finding notes and annotations in LiquidText at 4.4 out of 5. Since participants generally kept their excerpts and comments in the workspace, this suggests the system worked well for them. Rather, the reason again may be that these functions were simply not needed for the particular active reading task given to participants. For example, P17 pointed out that she did not need to pan/zoom the workspace (her comments/excerpts comfortably fit on one screen, as shown in Figure 7.15), but noted that she might for a longer task. This idea was echoed by P18, who likewise felt he might have had to use zooming, as well as the magnifying glasses, if the document had been longer. Thus, there is evidence that some of LiquidText's more advanced features may not have been appropriated by users, at least in part, simply because the task I selected did not require them.

7.3.3 Answering the Research Question

The fourth research question asks what impact LiquidText's representation had on the process of active reading, and to what extent its functionality was appropriated into people's reading processes. And to summarize by results above, the answers to both of these questions appear to center around the use of an intermediate document as part of the active reading process. That is, LiquidText tends to lead people to create a summary-like document in their workspace composed of excerpts and comments. These excerpts and comments provide a condensed, sort of digested, version of the original article containing the key ideas and points relevant to the reader's larger task. Participants then commensurately adjust their overall workflow such that the early part of their reading process includes the creation of the intermediate document, and the latter part includes navigating and reading it. Consequently, the annotations/notes readers create tend to focus more on summarization than analysis, with readers apparently performing critical thinking and analysis in their minds without encoding them in the reading medium.

But while LiquidText promoted this body of process changes, my concern, as stated in the research question, is what changes were caused by the flexible, high degree-of-freedom representation used by LiquidText, as opposed to idiosyncrasies of its implementation. To begin to disentangle these two factors, we can consider which aspects of LiquidText appeared to have the strongest impact on these behavioral changes, and these appear to be the functions most closely associated with the intermediate document workflow that participants employed while using LiquidText.

Perhaps the area of LiquidText most essential to enabling the creation and review of an intermediate document is the workspace, and considering the aspects of the workspace relied upon by participants is telling. First though, the very existence of the workspace was significant in providing flexibility, as it offered a way to organize and arrange materials without the risk of occluding or otherwise interfering with the original document. And by generally moving their materials to the workspace, participants appeared frequently to take advantage of this support. Beyond that, the workspace offered flexibility in arrangement, allowing participants to position and scale their materials freely, without forcing it into a predefined structure such as a list or outline. Participants took advantage of this as well by employing widely varying arrangement strategies as required for their preferred means of performing the task. Of course the workspace offered other features that could be characterized as offering flexibility, such as the magnifying glasses, but participants did not use these as regularly and so they likely had less impact on participants' reading processes.

Going hand-in-hand with the workspace, excerpts likewise offered a way for participants to leverage LiquidText's flexible document representation. First, excerpts effectively acted as a way to break up a document and see only the desired pieces. Furthermore, it provided a "safe" means of doing this, as the document pieces, in this case, were always linked to their sources so the risk of losing context was abated. The result was that users had a more flexible way to view their text, effectively allowing them to easily and safely construct a customized, selective view of their document. And indeed, this was how many participants employed excerpts, as with P13 and P18 who used them to make comparisons, and P15 who used them to organize parts of the text into groups. Likewise, the fact that excerpts can be connected provided another degree of control, as users could express relationships not only by links and spatial proximity, but by visual connectedness—a function used by 7 of the 12 participants. But more broadly, this facility for creating such a custom visualization appears to have been integral to participants' creation of an intermediate document, as the free arrangement of collections of excerpts was their primary tool for doing so. Thus, excerpts provide a prime example of how representational flexibility affected the active reading process.

And comments provide another such example. Specifically, and unlike in other media, users were able to remove their comments from the referent document and organize and arrange them in the workspace. This organizational flexibility was used extensively, with comments in the workspace outnumbering comments attached to the document by a factor of 2.8. And this flexibility was critical in enabling the intermediate document workflow, because it allowed users to review their comments without having to consult, or even look at, the original document.

Cumulatively then, I argue that LiquidText's impact on participants' behaviors was not a result of idiosyncrasies of its implementation. Rather, the major changes in participants' reading processes followed from aspects of LiquidText directly related to the flexible, high degree-of-freedom representation I sought to embody. Thus, the intermediate-document centered workflow can be seen as the impact of LiquidText's type of document representation.

And likewise, we see that the features of LiquidText which were most appropriated tended to be those most necessary for this intermediate-document workflow. Thus, functions relating to a simple linear reading of the original document, and the populating and navigation of the workspace, were among those most extensively relied upon by participants. By contrast, functions providing more complex, non-linear navigation of the original document tended to receive less use in participants' reading processes. But an application's functions are not used in isolation—the set of features

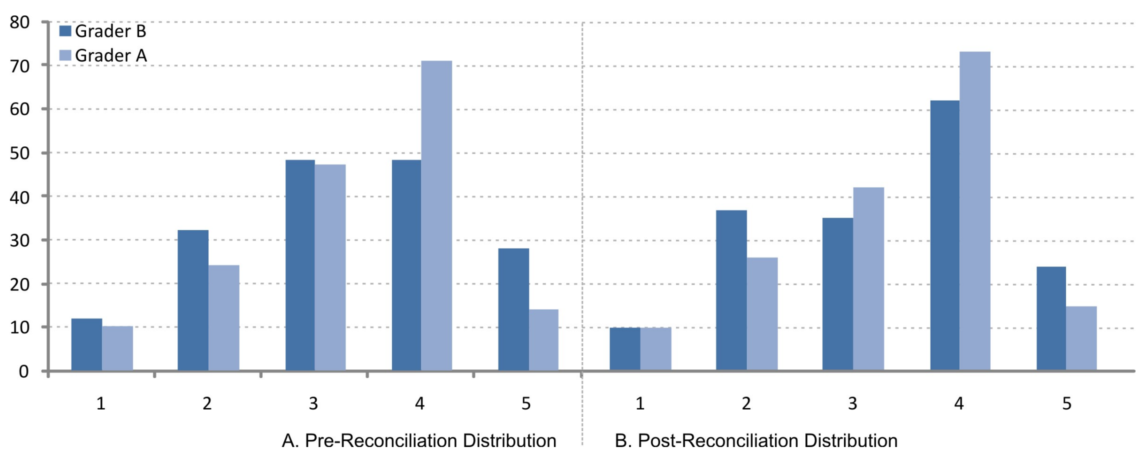


Figure 7.16. Essay grade distributions (averaged across all grading dimensions and experimental condtions) a) before the graders reconciled their scores, and b) after.

which were appropriated or not appropriated was closely related to the task at hand. And while I did work to create a relatively generic active reading task, various limitations, especially the time each participant could afford for the study, did limit the breadth of the task. As such, some features that were not appropriated may have proven useful with a longer active reading task. As participants discussed, this might include some of the more powerful workspace navigation functions, such as the workspace zoom gesture, and the workspace magnifying glasses.

7.4 Outputs of Active Reading

The basis for understanding the impact of LiquidText's representation on the *outputs* of the active reading process is, of course, through the essays which participants wrote. As explained in the analysis section, the essays were scored by a pair of graders along seven dimensions, who subsequently consulted with one another to reconcile their scores where possible (see Figure 7.16 for overall score distributions). I assessed the agreement between the graders using Pearson's Product-Moment Correlation, one of several accepted techniques for measuring inter-rater agreement [Bordens and Abbot 2008]. I chose this approach over the more traditional Cohen's Kappa because the latter neglected that grading was performed along a continuous scale, and thus treated very close scores as completely different. I could have accommodated this through bucketing, but Pearson was a more natural approach. Thus, I found that, even before reconciliation, the correlation coefficient $r = 0.78$, indicating that the scores are significantly correlated at $\alpha = 0.01$. Notably, the means/standard deviations of the scores were also well matched (3.14/1.04 and 3.17/1.12 for the graders respectively). After reconciliation, the inter-rater agreement was even stronger with Pearson $r = 0.93$ (means/standard deviations were 3.17/1.05 and 3.15/1.09 for the graders respectively).

Apart from the graders themselves, there are also two other notable factors that did *not* have a significant effect on the scores. The first is the order in which the conditions were completed. That is, essays written in the conditions that were run first for a given participant actually scored 9.7% higher than the essays in the conditions run second. Predictably, this small difference is not significant (single factor ANOVA, p = 0.60), and so we can likely dismiss concerns of learning effects between conditions. The second factor is which of the two articles the participant was to critique. As with order, there was no significant effect of the article on essay score, with one article associated with essay scores only 1.4% higher than the other (single factor ANOVA, p = 0.94). As such, we can generally assume it very unlikely that condition order or article had an effect on essay scores.

Throughout the rest of this section, I first discuss the ratings of these essays in more detail and consider where participants were more or less successful. I subsequently build on this by consulting the interviews and questionnaires, using participant feedback to help elucidate the results of the grading. Finally, I provide an overall concluding discussion of LiquidText's impact on active reading outputs.

7.4.1 Overall Ratings

On the whole the results of the essay grades for the two conditions present uncertain conclusions. Of the seven dimensions along which the essays were scored, six

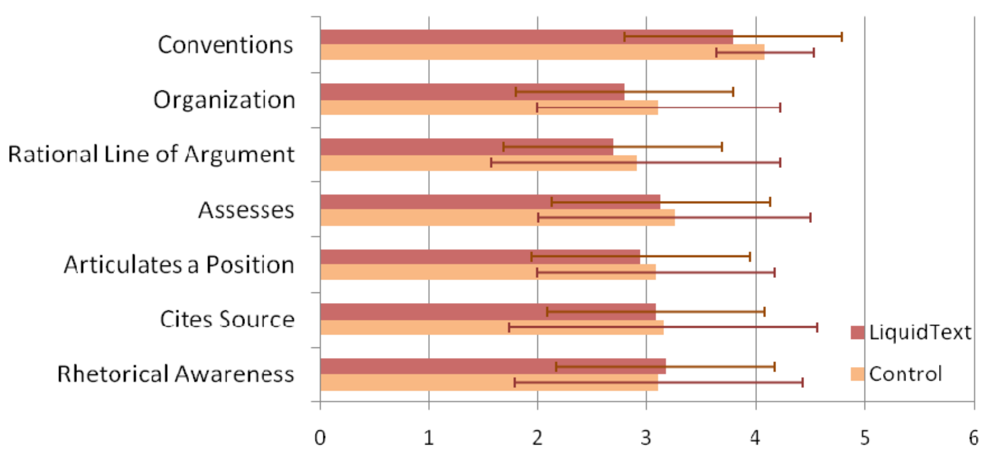


Figure 7.17. Average essay scores for the seven essay grading criteria for participants while using LiquidText and while using the control. Error bars depict standard deviation.

were slightly higher with the control condition, and one with LiquidText. But these small differences were not remotely significant at the $\alpha = 0.05$, or even $\alpha = 0.1$, levels. Thus, I see no consistent, generalizable effect of LiquidText on the outputs of active reading (Figure 7.17).

7.4.2 Scores by Participant

To explore why LiquidText had no significant net impact, we can look at the scores in terms of participants. One basic observation about participants' scores along the seven dimensions of the grading is that they are highly correlated—so participants that did well on "Rhetorical Awareness" also tended to do well on "Cites Sources." Indeed, scores for "Articulates a Position," "Rhetorical Awareness," "Cites Sources," "Assesses," and "Rational Argument," all had Pearson $r > 0.90$. Correlations involving "Organization" and "Conventions" were lower, but still had $r > 0.62$. Thus, rather than consider each component of the grading separately, I will principally consider the mean score for each participant.

Thus, looking at participants' scores broadly, we find that five participants scored

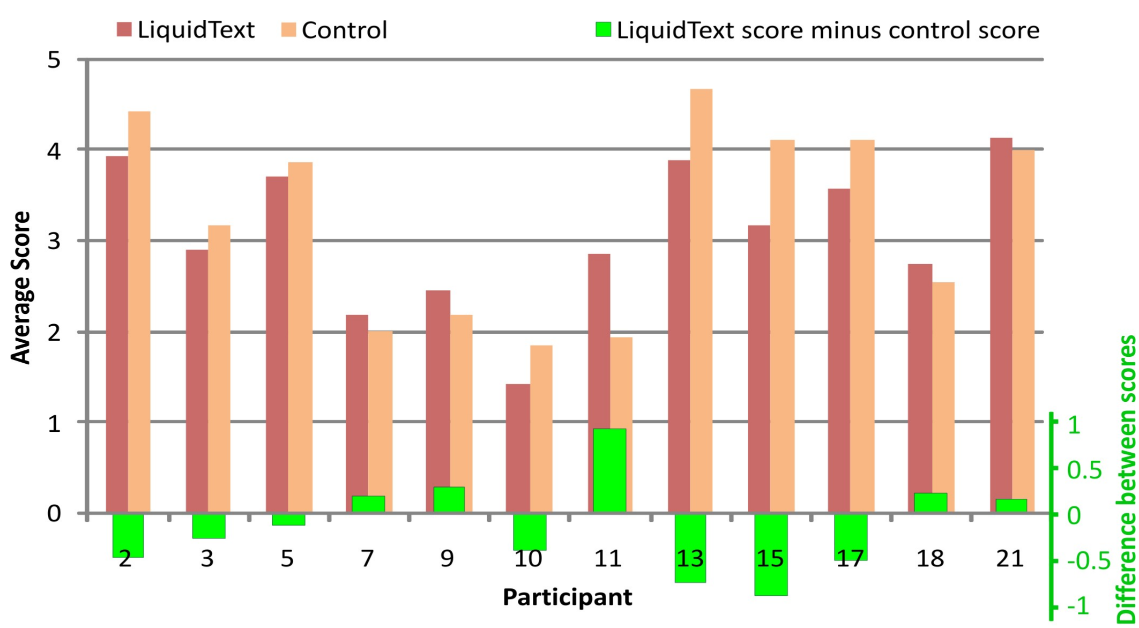


Figure 7.18. Average score for each condition for each participant. Also shows difference between the scores (in green), with positive values corresponding to a higher score with LiquidText.

better on average with LiquidText, and seven scored better with the control. And while participants' score differences are too small to conclude that LiquidText either helps or hurts active reading performance on the whole, we can consider the differences between those people whose scores improved with LiquidText and those whose scores did not. In this section I consider some factors that may help differentiate these two groups in order to better understand what led LiquidText to have the effect that it did.

To do this, I considered correlations with *score difference*: I use this term to refer to a participant's average score with LiquidText, minus their score average score with the control. In performing this analysis, I found a litany of factors that did *not* correlate with score difference. Factors such as gender, number of diary entries, and various components of cognitive load had no significant correlation with score difference. Neither did the extent to which a participant liked LiquidText, how enjoyable they thought it was, or how many annotations they made using it. However, even a cursory glance at Figure 7.18 suggests that the higher performing participants tended to do better in the control condition. And this effect was significant: for a given participant, mean essay score was weakly but significantly inversely correlated with score difference (again, LiquidText essay score minus control essay score) (Pearson $r = -0.51$, $\alpha = 0.05$). In other words, the participants who improved with LiquidText tended to do worse overall. The trend of struggling readers performing better with LiquidText is reflected in several other areas as well. There was a strong and significant inverse correlation between perceived enjoyableness of the article and score difference (Pearson $r = -0.70$, $\alpha = 0.05$). Similarly, participants who made fewer excerpts during the active reading task likewise tended to have done significantly better with LiquidText compared to the control (Pearson $r = -0.67$, $\alpha = 0.05$). Perhaps most surprisingly, participants who were not native English speakers tended to have a bigger improvement with LiquidText versus the control (either more positive or less negative), as compared to their native-speaking counterparts (ANOVA, $p < 0.01$). Note though, that among just the six non-native speakers, the difference between the LiquidText and control essay scores was not significant—so we cannot conclusively say that LiquidText helps non-native speakers, only that it affects them differently.

To help understand why participants did not perform better with LiquidText, it seems possible to look to the process differences discussed in the previous section. Indeed, one is tempted to speculate that the focus on an intermediate document, and thus summarization rather than analysis, led participants to underperform. But while such a scenario seems possible, it is also unlikely. The extent to which a participant created fewer analytical comments in LiquidText versus the control was not significantly correlated with essay score difference (Pearson $r = 0.15$). That is, just because a participant made fewer reflective comments in LiquidText than in the control did not mean she would also write a worse essay than in the control. Similarly, other ways that participants' reading processes varied between conditions, such as differences in the numbers of non-reflective comments and non-textual annotations, were likewise nowhere near significantly correlated with essay score differences. Similarly, the timelines, as in Figure 7.11, do not appear to vary in any predictable way between those for whom LiquidText raises versus lowers scores.

So if the focus on an intermediate document does not explain LiquidText's impact on active reading outputs, let us again consider what we do know. Summarizing the above, it appears that the participants most likely to improve with LiquidText tended to a) be poorer active readers, b) make fewer excerpts, and c) be non-native speakers. One possible interpretation of these data is that LiquidText's atypical affordances help those who are struggling active readers, but act as a disruption for those who already possess highly refined active reading skills. That is, even if LiquidText's functions are potentially beneficial, the disruption some of them present, at least in the short term, outweighs their potential benefit to refined readers. If true, this may explain why we do not see differences between the LiquidText reading processes of the two groups: the difference may not be in how the groups used LiquidText, but how expert they were at active reading with the control. And while evaluating this hypothesis is beyond the scope of this study, it does present a possible direction for future work.

And while LiquidText alone does not in itself have a significant overall impact on essay scores, other data did help explain how well participants' performed on their essays—especially the diaries. I found that there was a significant, negative correlation between the number of times participants recorded creating embellishments on documents in their diaries, and their essay score (Pearson $r = -0.74$, $\alpha = 0.01$). That is, the more often participants created embellishments, the worse on average they performed (judged by taking the average score across the two conditions). Likewise, using a multiple regression analysis including all of the diary statistics (that is, how often a participant created embellishments, produced annotations, etc. as recorded in their diary), I found that the diary data were able to capture 75% of the variance in a participant's average essay score (i.e., $r^2 = 0.75$). The number rises past 80% when we replace average essay scores with *just* essay scores from LiquidText or just essay scores from the control ($r^2 = 0.82$ or 0.88, respectively). Thus, as one would expect, this suggests that most of the variation in participants' scores was related to their active reading experience rather than the media employed.

7.4.3 Participant Perceptions of Outputs

While participants' backgrounds and reading expertise may correlate with their essay scores, and LiquidText's impact on their essay scores, their beliefs about their performance generally do not. We see this strongly in the interviews, which failed to elucidate any clear attributes associated with participants who excelled with LiquidText versus those who did not. In particular, participant' reflections about their own performance seemed largely unrelated to their actual performance. P17, for example, strongly preferred LiquidText, and said, "If you are not comfortable with the task or do not want to do a task you are not going to do it as well," in explaining why she felt she performed the task better with LiquidText. Yet she scored over 17% higher in the control condition. P15 likewise felt that he would be more effective with LiquidText, yet performed over 43% better in the control condition. Conversely, P7 generally appeared to favor the control, saying that it takes more time to deal with the type of environment made available with LiquidText, yet he performed 18% better in the experimental condition.

Many of the questionnaire responses likewise bear this inconsistency. Recall from section 7.2 that participants generally found LiquidText to provide a significantly better reading experience than the control. As shown in Figure 7.1, participants felt their thinking process was better with LiquidText, and that it was more enjoyable and overall preferred. Most strikingly, as shown in Figure 7.1 Q15, they felt on average that their responses were overall better with LiquidText as well. Yet not only were participants' LiquidText scores not higher overall, but there was virtually no correlation between where participants felt they wrote better responses and where they actually did. This can be seen in two ways: First, we can consider the correlation between essay score difference (again, the participant's overall score on LiquidText minus their overall score on the control) and the survey question directly asking where participants felt they wrote a better response (in which higher values favor LiquidText and lower favor the control). The correlation was negligible, with a Pearson $r = 0.2$. Second, we can consider essay score difference versus the difference between participants' TLX performance measures using LiquidText and the control. Again, there was no correlation, with Pearson $r = 0.08$. Replacing Performance with the other TLX measures yielded similarly negligible correlations. Finally, we can consider the simple correlation between essay score and the TLX performance measure, as well as a survey question inquiring how effectively the participant thought they read the passage. These revealed correlation coefficients of $r = -0.06$ and 0.17 respectively (see Figure 7.19 for a summary of these correlations).

Cumulatively then, we see that participants misjudged not only the absolute ratings of their essays, but also the relative differences between their essay scores across the two conditions.

7.4.4 Answering the Research Question

The fifth and final research question asks what impact LiquidText's flexible representation had on the outputs of active reading. And most generally, it appears that it did not have any significant impact at all. This is somewhat paradoxical for two reasons. First, in spite of their actual performance, according to many measures, participants believed that they were writing better essays with LiquidText (e.g., Figure 7.1 Q15). And according to all measures, participants significantly and consistently preferred LiquidText to the control media. The second paradox is that in spite of the negligible effect on essay score, LiquidText significantly affected participants' reading processes. It is surprising that the vast changes described in the previous section did not generalizably affect essay scores.

But while my data do not provide a full resolution to these paradoxes, part of understanding them requires the recognition that LiquidText does not have a singular impact on the outputs of active reading. Different readers' outputs are affected in different ways, especially depending on their reading ability as discussed above. When this is factored in, LiquidText's effect on reading outputs, perhaps owing to its impact on

Variable 1	**Variable 2**	**Pearson Correlation Coefficient**
LT essay score – ctrl essay score	Survey question asking where a better response was written (higher favors LT)	0.2
LT essay score – ctrl essay score	TLX performance rating on LT – TLX rating on ctrl	0.08
Essay score	TLX performance rating	-0.06
Essay score	Survey question asking about effectiveness at reading passage	0.17

Figure 7.19. Summary of correlations between participants' perceptions of their essay scores and their actual scores.

the reading process, can be seen more clearly. For example, the TLX measures do not explain score difference very well on their own—however, a multiple regression including the TLX scores, as well as the control condition essay scores (used to represent baseline active reading ability), explains over 89% of the variance in score difference[10]. Similarly, the differences in the types of annotations/notes created in LiquidText versus the control do not in themselves explain score difference. But when control condition essay scores are included, they can together explain 77% of the variance in score difference[11]. What these statistics mean is that differences in the cognitive demand (reflected in the TLX scores) and differences in reading process (reflected in the numbers and types of notes/annotations created) do help explain LiquidText's impact on participants' active reading outputs—but only when we take into account that these impacts vary depending on the expertise of the reader.

But while the data I gathered enables us to see hints of the complex interplay between personal background and LiquidText's functionality on active reading outputs, they also point toward the type of future studies that will enable us to gain a clearer, more complete picture. Ultimately, to understand and disentangle the effect of reading skills, native language and other factors, my future work would need instruments that probe these issues directly. Likely, this would include a variety of reading assessment tests, more nuanced interviews about participants' longer term active reading background and

---

[10] For this, I conducted a multiple regression analysis using six variables to capture the variance of score difference. The six variables were 1) control condition essay score, followed by 2) TLX mental demand on LiquidText minus TLX mental demand for the control, and so on for TLX's physical demand, temporal demand, and performance measures.

[11] Here, I conducted a multiple regression analysis. The input variables were 1) control essay scores, 2) reflective comments in LiquidText minus reflective comments in control, and likewise for excerpts and non-reflective comments. The variable being modeled was, again, score difference.

history, and substantially larger numbers of participants, allowing for higher levels of statistical significance for small subgroups of the sample.

But as such, we do not yet have a conclusive answer to the question of how LiquidText's flexible, high degree-of-freedom representation affects the outputs of active reading. While one can make some claims about LiquidText's impact, disentangling the effect of the representational approach from that of the implementation would require a deep understanding not only of what, but of *how*, LiquidText affected reading. Thus, the answer to this research question points toward future work where we can more completely address it.

7.5 Revisiting the Thesis Statement

As I discussed in the introduction, the objective of this dissertation has been the investigation of my thesis statement:

> *Active readers require more spatial flexibility than is available in paper or present computer-mediated reading environments. Therefore, giving readers a computer mediated environment with a high degree-of-freedom visual, spatial representation, along with a comparably high-dimensional input method, will result in a subjective and objective improvement in the active reading process*

The first part of the thesis—that readers require more flexibility—was addressed in Chapter 3. And now, from the findings revealed throughout this chapter, we are in a position to address the latter part of the statement—that giving readers a high degree-of-freedom representation and input method would result in subjective and objective improvements in active reading.

7.5.1 The Subjective Component

That LiquidText's representation results in subjective improvements in active reading seems virtually certain. As I explored in section 7.2, participants consistently preferred LiquidText overall, and likewise preferred it for most of the processes within active reading as well. Particularly, tasks such as finding notes and annotations, working with different parts of the text at once, and keeping track of relationships within the text were all seen as considerably easier in LiquidText. Of course, LiquidText also detracted from the reading experience in some ways, resulting in higher levels of cognitive difficulty on several dimensions of the TLX load text. Significantly though, the interviews suggested that the aspects of LiquidText leading to these difficulties centered on details of its implementation and hardware, including the text selection mechanism and the size of the touch screen. By contrast most of the functions, and broader attributes of the system, that participants praised tended to be examples of the flexible document representation LiquidText was intended to embody.

As I argue in Chapters 1 and 4, when giving users a representation with many degrees of control, it is important to provide a commensurately high dimensional input device—in this case, a multitouch display. And while it is challenging to disentangle participants' reactions to the input modality from the interactions, participants generally reacted well to the touch-based interface. Simply recalling and performing the gestures tended to pose little problem, as participants rated the ease of doing so at 4.6 and 4.2 out of 5, respectively. Likewise, of the six participants who addressed LiquidText's use of multitouch in the interviews, five reacted positively—with the exclusion of text selection. P7, for example, was not supportive of LiquidText's use of touch, and felt that a mouse would have made the task easier, especially for writing. However, he was in the minority; P13, by contrast, felt the type of task we assigned was well suited to touch, and P18 felt it was easier to use one's finger than a traditional mouse and keyboard combination. P17 put it concisely, saying, "No, multitouch is definitely the way to go." The key point then, is that my use of a high dimensional input modality generally felt consistent with, and did not detract from, the experience of using LiquidText. As such, I believe we can reasonably conclude that, *giving readers a computer mediated environment with a high* *degree-of-freedom visual, spatial representation, along with a comparably high-dimensional input method, results in a* <u>*subjective*</u> *improvement in the active reading process.*

7.5.2 The Objective Component

But as explained in sections 7.3 and 7.4, the situation for the objective impact on the active reading process was more difficult to assess. The impact on the reading process itself was, of course, substantial, leading participants to the use of excerpts and the LiquidText workspace to create an intermediate document. And while comments were still frequently created, they were used more for summarization and heading, rather than playing the more analytical role they had in the control condition. Indeed, generally, LiquidText's tools and approach to representing documents seemed to encourage less written analysis during the reading process than the control condition. But whether these changes constitute an improvement or a detriment is difficult to ascertain. While participants did produce fewer analytical comments in LiquidText, we cannot say for sure whether participants were performing less analysis, or simply performing it in their heads, as P15 had suggested. Notably, the decrease in analytical comments was not correlated with essay score difference.

Cumulatively, though, participants' objective performance using LiquidText was generally not better than their performance in the control. Indeed, the 12 participants in the study on average received scores slightly lower using LiquidText than the control—though this difference was too small to be statistically significant. However, a deeper exploration of the data showed that participants with lower levels of active reading expertise were more likely to benefit from LiquidText than those with greater expertise. While the cause and extent of this effect is still uncertain, it does point toward a more complex, reader-dependent effect of LiquidText on the reading process. As such, I do not believe that we can either conclusively assert or deny the last aspect of my thesis statement, namely that: *giving readers a computer mediated environment with a high degree-of-freedom visual, spatial representation, along with a comparably high-dimensional input method, results in a* <u>*objective*</u> *improvement in the active reading process*. Ultimately, the answer to this question will require further investigation, which I discuss in the following chapter.

7.6 Limitations of the Study

As discussed in Chapter 6, the summative study was designed to assess active reading in a representative, general way. But as with any lab assessment, there are practical constraints on the design of the study—and these constraints can lead to limitations on the generalizability of its results. And in the case of a phenomenon as vast as active reading, nearly any study will have multitudes of constraints which limit its generality. Thus in this section I consider several of the most noteworthy of these constraints and their potential impact on the findings presented in this chapter.

In this summative study, many of the limitations on generalizability came from the nature of the reading assessment task itself. Among these, an especially noteworthy factor was the length and difficulty of the task. Due to practical constraints on study duration, participants had only 55 minutes for the reading task in each condition; as such, the task itself was necessarily designed to be performed in approximately 55 minutes, including both reading the original article and composing a response essay. The above findings thus must be seen as representing a relatively short task with relatively short documents.

Having performed the study with a longer task or document could have resulted in several possible differences. The most obvious of these is feature use. In LiquidText, for example, participants pointed to several advanced functions—such as the magnifying glasses and workspace zooming—that they felt might have been used more often had the document or task been longer. Additionally, a longer task with longer documents might have required participants to integrate a larger volume of material in creating their critiques, as well as simply maintain awareness of more document content. In such cases, LiquidText's centralized workspace might have proven to be a more valuable asset in relieving cognitive load.

A similar constraint comes from the time-distribution of work on the task. Real world active reading tasks can be spread out over days, weeks, or months—such as a student who reads a textbook and then reviews their notes only before the final exam. But due to time constraints, it was not practical to require participants to spend additional days attending the study. Instead, the entire task for each condition was performed in one sitting. Had the task been spread over several days, it is hard to estimate the effects on participants' reading processes with LiquidText, but it is likely they would be required to refresh their memories of the reading task at each session. In such a case, the intermediate document might serve to help readers reacquaint themselves with a text after having been away from it, perhaps adding an advantage to performing the task with LiquidText.

There are also broader limitations arising from the time constraints of the study. By the time participants completed the experimental condition, they had only used LiquidText for under three hours. In contrast to years or decades of experience using the control media, participants had relatively little time to learn how best to appropriate the functions offered by LiquidText. Additional time could have led participants to better familiarize themselves with the system, and possibly incorporate a wider range of its functions into their reading process. For example, P15 even described that he forgot about a function that would have been useful, even though he could perform it during the training. More time would also have provided a better opportunity to develop a reading process optimized around LiquidText's features. That is, although participants did change their process with LiquidText, they might have made more refinements and further optimized their process if given more practice. Cumulatively then, it is important to see the above results as representing LiquidText's impact on reading relatively soon after adoption, as they may change considerably over more time.

Like the distribution of work across time, the task was also limited in its distribution of work across people. As I discuss in Chapter 3, active reading tasks often involve collaboration, which sometimes includes multiple people working with the same document. By contrast, the results of the summative study must be seen as representing decidedly solitary active reading work. Had the task included multiple people, either in synchronous or asynchronous collaboration, readers' processes could have changed in various ways to include inter-participant communication. One possibility is that the workspace might have acted as a common space for participants to share their reactions to the text, in which case they might have created more analytical comments in order to obtain feedback on their thoughts from the other participants. But this is just one example—adding multiple participants would have raised numerous issues, such as negotiating how resources like the workspace would be used, how work would be apportioned, etc.

Beyond the study's design, the sample of the population who participated was also constrained. As discussed in Chapter 3, active reading is performed by people in many contexts and roles, from students to designers to executives. But in the summative study, it was not feasible to recruit a sample that was truly representative of the active reading population as a whole. Rather, all twelve participants were graduate and undergraduate college students. Although the students did represent areas of study ranging from business to human centered computing, it is still not clear that they encompassed the same level of diversity that would be found in the active reading population as a whole.

Still, the use of a relatively homogeneous sample for the summative study does not render these results irrelevant apart from college students. As the findings of Chapter 3 show, the problems and preferences of active readers do occur commonly even across a highly diverse group. Rather, these results should be taken as starting points for understanding what questions to ask and how to formulate broader studies that target wider, or simply different, samples of the active reading population.

Generally then, the summative study represented a specific example of active reading, and assessed LiquidText's impact thereupon. The study was designed to incorporate many attributes typical of active reading tasks—such as working individually, critical thinking, reading to create a written output, and time constraints—but none of these attributes are universally present in all real world activity. Rather, active reading is sufficiently broad that no singular experiment could likely fully capture its breadth. And while the use of a lab study does allow for control, it can exacerbate this issue by potentially divorcing the active reading task from the types of larger circumstances that surround reading in the real world—such as long term deadlines, collaborators, etc. As such, this study should be seen as providing a starting point to understanding the impact of a flexible, high degree-of-freedom representation on the reading process.

7.6.1 Behavioral Limitations

In part as a result of the constraints of the study, the behaviors demonstrated by participants in performing their tasks do not encompass the full breadth of behaviors seen in the literature. While many such differences could potentially be identified, in this section, I briefly review some of the more prominent examples..

One approach to elucidating which participant behaviors were absent relative to typical active reading practice is to consider the goals which participants were pursuing. That is, we would not expect people to behave the same way when proofreading as when writing a critique. Thus, we can begin by observing which of the reading goals discussed in Chapter 2 were, or were not, present in this task. Most obviously, tasks like *proofreading* were not present, since participants were not asked to consider minor grammar/spelling issues in their critique, and correspondingly did not appear to do so. *Reading to learn*, a general investigation of a subject, was likewise absent, as was *reading while writing from multiple sources*, and *reading to support discussion*. Since participants were not operating with these goals, it was not surprising to find that they did not employ many of the behaviors associated with these goals. For example, many of these goals involve reading multiple source documents in parallel—such as those involving cross-referencing, or *reading for research*. Thus, participants did not have to integrate content from multiple documents, nor did they have to ascertain a sense of the space of possible documents one could read—things that can be an important aspects of other active reading tasks. Likewise, the brief, skim-reading sometimes used in *reading to learn* was also absent in this case.

Looking more broadly across reading goals, we can consider participants' behaviors in terms of some of the major categories of active reading phenomena discussed in Chapter 2, including navigation, annotation, and excerpting. In the context of navigation, one notable idiosyncrasy of the summative sudy is the general linearity with which participants read the original document. As discussed above, participants typically read through the text linearly, creating excerpts or comments, or highlighting and underlining, before proceeding on to write their critique. But active reading often entails extensive non-linearity as readers zig-zag through a text searching for specific content, or jump ahead looking for previews of what is to come. While readers in this study did exhibit some of these behaviors, they generally did so later in the task, after they had already read the document through completely. Likewise, due to the short length of the text, some of the more elaborate bimanual interactions were also absent—like holding a finger in a particular page of a book while flipping through the remainder. As discussed in Chapter 5, LiquidText's bimanual counterparts to these interactions were likewise rarely used.

In contrast, participants' annotation behavior broadly encompassed more of what we find in the reading literature. Participants created a wide variety of markings, as well as margin notes and notes on separate surfaces (although to a lesser extent in the control condition, as discussed above). Still, annotation practices are known to vary depending on whether the document is intended to be shared [Qayyum 2005], so by including only individual tasks, this study necessarily did not capture the full breadth of annotation practice.

Similarly to annotation, many of the behaviors typical when writing on a secondary surface were indeed exhibited by participants in the summative study. However, some limitations were also present—the most striking being the general *lack* of writing on a secondary surface in the control condition. More narrowly though, we can consider the specific types of information readers tend to extract to secondary surfaces. For example, participants generally did not extract bibliographic information [O'Hara, Smith et al. 1998], which was of course to be expected, as there was only one document to be read. The reasons for which people used a secondary surface were also somewhat more limited, as other research has shown people sometimes use their notes to act as a more portable version of the original text [O'Hara, Smith et al. 1998]. But in this case of course, portability was not an issue since participants remained in the same location. Likewise, people sometimes take notes on separate surfaces as a means of integrating content from multiple sources—but in this task of course, there was only one source and so such integration was not necessary.

The notable absence of portability is also in contrast to my own studies (Chapter 3). There, people were found to enjoy taking their reading material to various locations ranging from libraries to kitchen tables. In this case though, we prevented that behavior by requiring them to perform their tasks at desks in a lab. Also in contrast to my earlier findings, participants performed their task in isolation of a complete active reading workflow. Thus common active reading behaviors, such as handing a document off to others for review or discussion, were absent.

Cumulatively then, participant behavior was substantially narrower than what we would find in the reading literature. However, this narrowness generally comes from the task which the participants were assigned. Within the context of such a specific, controlled task, participant behavior, particularly in the control condition, was largely as one would expect. Still, there was a surprise in the general lack of writing to a secondary surface during the control condition. And while we do have some understanding of why participants *did* use the secondary surface in LiquidText, the exact reasons they did not do so in the control condition are less certain. One possibility is that the brevity of the task did not appear to demand it, and participants may have found it easier to take notes alongside their text in the control. Ultimately though, I leave this to future work investigating LiquidText's impact on tasks of greater duration, which will likewise elucidate how the behavior elicited by a longer control condition varies from expected reading behavior.

CHAPTER VIII
CONCLUSION AND FUTURE WORK

This, the final chapter of my dissertation, is organized into two sections: a recap of the discussion so far, and future work. The first provides a concise summary of the major findings of this research, as well as the motivation and design of LiquidText. It also offers an opportunity to review the answers to the research questions in context. The second section summarizes the various opportunities discussed in this dissertation for continuing and building upon this research. I discuss each in turn.

8.1 Recap of Dissertation

The core goal of this research has been to investigate a new approach to document representation in order to better support active reading. But underlying this is the observation that active reading offers an opportunity to be improved. I investigated this opportunity through the active reading study described in Chapter 3, which involved diary studies, interviews about the reading experience, and a design workshop. The study provided two areas of findings—the first being an update and refinement to earlier understandings of active reading. For example, the study helped to shed more light on personal active reading, as well as on the breadth of places and circumstances in which active reading occurs. It showed the increasing prevalence of computer-mediated reading, and confirmed both the necessity and the difficulty of using computers and paper in parallel. More interestingly, it showed the variety of materials used for input to, and output from, the reading process—especially the prevalence of email.

Most importantly though, the study helped show the problems people face with active reading. These included problems with visualization, such as seeing the desired amount of text or comparing different parts of a bound document. Many problems also related to annotation, as participants described challenges of running out of margin space, and the difficulty of finding annotations after creating them. Similarly to annotation, participants also pointed to challenges expressing and maintaining awareness of relationships that happen to span multiple documents. Navigation posed possible problems as well, as readers struggled to switch among their documents, even including notes. Other difficulties included the maintaining of awareness of relevant background information during reading, as well as issues creating and organizing notes. Together, these and other problems pointed toward a reading experience that is in many ways inflexible, where it is difficult to arrange, visualize, organize, and navigate one's reading materials with the freedom and control required for the task at hand.

These findings painted a rich picture of the state of active reading, and they form one of the central contributions of this research—updating our understanding of the active reading process, and exposing the difficulties readers face. These findings also address my first two research questions, respectively.

Rather than address the problems of active reading in a piecemeal, isolated way, I sought a broad design approach in an attempt to offer a consistent, complete alternative to existing reading media. That approach focused on providing more flexibility and fine-grained control over the visual arrangement, navigation, and relationships among content. And to investigate this approach, I designed and developed LiquidText, a multitouch active reading environment that emerged from this design direction, as well as the results of my and others' earlier reading studies.

In order to develop LiquidText from an abstract concept to a useful, usable embodiment, I employed a multi-step iterative design process. After developing a basic prototype informed by prior reading research, discussed in Chapter 2, I conducted a formative lab study in which participants provided feedback after using LiquidText to perform an actual active reading task. This feedback led to a host of feature additions, including the workspace, fisheye magnifying glasses, highlight aggregation, and linking, as well as various improvements to the existing functions. I evaluated this refined system principally in terms of its gesture interface. This included weeks of rapid iterative testing in which I repeatedly evaluated users' abilities to learn and perform LiquidText's gestures, followed by changing those gestures and repeating the cycle several days later. The final system had substantially refined interactions for several functions, including collapse, highlight aggregation, text selection, and others.

Many of the core contributions of this research came out of the design and development of this LiquidText system. In part these consisted of elements of the LiquidText user interface. A variety of interactions, including collapsing, multitouch magnifying glasses, and others could be applied directly to other applications or application domains. The same is true of comments that can pertain to arbitrary numbers of passages, be organized and interconnected independent of the source document, and maintain their link to their referents.

To evaluate the LiquidText system, and the design approach that inspired it, I conducted a summative evaluation. The goals of the evaluation were to understand the impact of a flexible document representation, like that used by LiquidText, on the subjective experience, process, and performance, of active reading. To investigate this, I ran a controlled lab study including twelve participants who performed a 55 minute active reading task using LiquidText as well as a more traditional medium. Over the course of the study, participants completed various questionnaires and interviews, spent two days learning LiquidText, and recorded a reading diary before the start of the lab sessions.

The results of the summative evaluation found that users responded positively to LiquidText. They strongly preferred it to the control condition media, and found it excelled especially in tasks like navigation, handling annotation, content arrangement, and organization. Thus, many of the aspects of active reading found to be problematic in my formative study were the areas where participants felt LiquidText was most beneficial. But in spite of these benefits, LiquidText also appeared to raise users' cognitive burdens and generally not to provide an easier reading experience. The apparent paradox of LiquidText being both preferred but leading to a greater cognitive load was not fully resolved, but may have been the result of difficulties with a small number of gestures and the hardware on which the system ran. Still, the most lauded aspects of the system were those giving the user more control and freedom to arrange, organize and connect their materials—aspects emerging directly from the underlying design goal of providing a flexible, high degree-of-freedom document representation.

This design approach then, with the resulting impact on subjective experience, is one of the core contributions of LiquidText, and has potential applications in a variety of other domains. But even within the broader domain of active reading, it could be applied well beyond the scope of protracted, critical text consumption that characterized the design of LiquidText and the summative study. Formative study participants showed that active reading occurs in contexts ranging from email to spreadsheets, and many of the challenges of inflexibility noted above are present in those contexts as well. Thus, solutions like those used in LiquidText might likewise be beneficial in future email clients and spreadsheet applications, among others.

But more broadly, we could also look outside active reading entirely, exploring other domains that involve tasks analogous to those of active reading. Areas that involve comparison, non-linear navigation, annotation, and so on may well be amenable to a LiquidText-like solution. Areas like Geographic Information Systems, video editing, or Computer Aided Design could be good candidates to merit exploration. Still, the particular mismatch between the requirements of the task and the capabilities of the media found in active reading may not be found to the same extent in other domains.

But beyond subjective experience, LiquidText's impact on active reading is more complex. In the case of the reading process, which is the focus of my fourth research question, when using LiquidText, readers tended to create and rely upon an intermediate *summary* document derived from original text, rather than the original text itself. Consistent with a focus on summarization, readers using LiquidText also tended to perform less written analysis in their comments and notes, as compared to readers in the control condition. And notably, the functions and attributes of LiquidText that enabled the creation of this intermediate document particularly included those derived from LiquidText's flexible, high degree-of-freedom design philosophy. Indeed, it appeared that participants proverbially jumped at the opportunity to restructure the visual presentation of their documents.

The impact of this change however, was less clear. For collaborative tasks, or tasks spread over multiple days, a rich summarization of a text could prove helpful for refreshing one's own memory or meditating communication with others. But for the short individual tasks used in the summative study, it could have led participants to dedicate too much time to summarization at the expense of analysis.

Ultimately though, even for the short, individual tasks used in the summative study, whether LiquidText's impact on the reading process was positive or negative—my fifth research question—was ambiguous. On average, LiquidText had no significant impact on the scores of participants' essays. But underlying this, individual participants were affected by LiquidText in very different ways—namely, the scores of just less than half improved with LiquidText, and just over half declined. The complete reasons for this were unclear, but broadly, those participants who were less advanced active readers, or non-native English speakers, tended to experience more of an improvement, or less of a decline, than others. But a full characterization of this disparity or its causes is beyond the scope of my data, thus a more complete answer to my fifth research question is left for future work.

8.2 Future Work

This research investigates the design, development, and evaluation of a new approach to supporting active reading. But in doing so, it raises a variety of questions that are necessarily beyond the scope of this dissertation. In this section, I discuss three areas of future work opened by LiquidText: 1) further evaluation, 2) the reading process, and 3) design. I discuss each of these in turn.

8.2.1 Evaluating LiquidText

While the summative study revealed much about participants' reactions to LiquidText, there are opportunities to evaluate and investigate the system in several additional ways. The first of these is to evaluate LiquidText in greater depth, further characterizing its impact on active reading performance. And perhaps the most salient approach to doing so is extending the study. While participants did largely report a high level of comfort with the LiquidText gestures, as discussed in Chapter 7, it was not obvious that they had developed the same level of refinement in their reading strategies as they had with the control media. More generally, there is an opportunity to vary many of the constants of the summative study, including the duration as well as the task itself. One could explore whether LiquidText's impact changes if multiple participants must collaborate during parts of the task, or perform a task that is extended over several days. Likewise, one could explore the use of different final products than essays, such as answering multiple-choice questions, or engaging in dialogs. Some of these might be helpful in disentangling LiquidText's impact on the performance of reading versus that of composition.

More critical though, is exploring why, even with the current task, participants' native language and active reading skill appear to have such a strong interaction with LiquidText's impact on reading performance. Exploring whether this effect holds when changing some of the above variables could shed light by elucidating the limits of the effect. Likewise, one could design studies to more carefully assess participants' reading backgrounds to better identify the ultimate cause behind this effect, rather than merely correlations.

But while we can understand the impact of LiquidText in greater depth through more elaborate lab studies, understanding its effect most broadly requires a different type of study entirely—a deployment. While a controlled, laboratory study is a critical element in elucidating the differences between LiquidText and existing media, a deployment offers the opportunity to observe how LiquidText fits into the real, authentic tasks and social milieu of active reading.

But as I explain in Chapter 6, a deployment includes many challenges, including the level of refinement of the artifact being deployed. LiquidText, though well developed for a research prototype, lacks even basic functions required for a real world tool. While some such features would be conceptually simple—such as saving and opening files—even basic support for collaboration, user pedagogy, and the like would require additional basic research and design before any deployment would be possible.

8.2.2 Reading Process

Beyond evaluating LiquidText's overall impact on reading performance, there are also opportunities to look more deeply into its effect on the reading process. One opportunity is to explore how the reading process if affected by the convenience, ease, and enjoyment of the different functions in the system. The summative study already hinted that some participants were more inclined to use the excerpting functions because they were perceived to be especially easy to use. One could thus explore different LiquidText designs, where different features were made more or less convenient, or given more or less engaging interactions, to observe how this causes them to be appropriated in active reading. While such research could be interesting in itself as elucidating the requirements and preferences of active reading, it also raises the potential for reading tutoring. As noted above, one could investigate whether such differences could help people become better readers by gently persuading them to use more effective reading processes.

8.2.3 Design

While the design of LiquidText involved several iterations of refinement, there were several additional avenues of investigation that were not feasible to explore. Some of these include simply improving existing elements of the interaction design. The fisheye magnifying glasses, for example, were sometimes difficult for users, leading to a sense of instability. In further iterations, one could evaluate alternate distortions, such as simple magnification, or a two dimensional version of the collapse visualization used elsewhere in the system.

Similarly important, there were several areas of functionality not included in LiquidText at all that participants suggested I add. One of these was a more flexible form of linking. That is, while LiquidText does allow strings of text to link, participants wanted to be able to link general areas of the workspace, or groups of objects, to one another. Supporting such breadth would likely be a considerable project on its own, requiring a largely redesigned linking interaction, as well as new visualizations to communicate the linking structure back to the user effectively. However, it would fit well with LiquidText's goals of giving the user more flexibility and control.

Additionally, even going back to the formative study, participants sought something like pen input in addition to, or instead of, touch. While most participants seemed comfortable with the use of a touch interface, some tasks, like text selection were challenging with fingers. Certain pen-oriented functions, like drawing, were entirely absent from the current system as well. Thus LiquidText offers an opportunity to explore the integration of pen and multitouch input in the context of active reading. And while such integration would likely be challenging, it could also raise the possibility of a rich array of new tools, as in [Hinckley, Yatani et al.].

[End]

Appendix A

Questions used in the diary portion of the formative study.

1. What was the larger activity you were doing?
2. What documents were you working with?
3. What medium/media did you use?
4. What physical setting did you do this in?
5. Did you embellish the documents you were reading?
6. Did you produce any other documents as a result of this reading?
7. Did you read straight through this document(s), or did you have to jump around in some way?
8. What were the hardest parts of the activity, especially in terms of the tools and media you used?
9. On the whole, what was your impression of performing this activity?
10. Can you think of any ways it could have been improved--you can suggest feasible as well as "impossible" ideas?

Appendix B

Instructions given participants for performing critique task in summative study.

We would like you to critique the article you will be presented with. We are not looking for an especially long or refined piece of writing from you, but we do ask that you try to think carefully about the article and look at its larger substance rather than minor issues of grammar and such.

In particular, try to answer the following in your critique:

- Overall, was the article good or bad?
- Was the article's argument sound?
- What points or ideas did you agree or disagree with (please be as specific as you can)?
- What, if anything, would you change about the article (again, please be as specific as possible)?
- Were there specific portions of the article that seemed contradictory?
- How strong were the supporting arguments? Why?
- How well connected were the supporting arguments?
- Were the supporting arguments relevant to the overall message?
- Did the speaker seem believable (i.e. non-biased, etc)? Why or why not.
- What type of supporting information was lacking in the paper that would have helped support the argument? You do not have to be very specific.

Appendix C

General rubric used by the Georgia Tech Writing and Communications Program.

Scale	1: Basic	2: Beginning	3: Developing	4: Competent	5: Mature	6: Exemplary
Rhetorical Awareness Response to the situation/assignment, considering elements such as purpose, audience, register, and context	Ignores two or more aspects of the situation and thus does not fulfill the task	Ignores at least one aspect of the situation and thus compromises effectiveness	Attempts to respond to all aspects of the situation, but the attempt is insufficient or inappropriate	Addresses the situation in a complete but perfunctory or predictable way	Addresses the situation completely, with unexpected insight	Addresses the situation in a complete, sophisticated manner that could advance professional discourse on the topic
Stance and Support Argument, evidence, and analysis	Involves an unspecified or confusing argument; lacks appropriate evidence	Makes an overly general argument; has weak or contradictory evidence	Lacks a unified argument; lacks significance ("so what?"); lacks sufficient analysis	Offers a unified, significant, and common position with predictable evidence and analysis	Offers a unified, distinct position with compelling evidence and analysis	Offers an inventive, expert-like position with precise and convincing evidence and analysis
Organization Structure and coherence, including elements such as introductions and conclusions as well as logical connections within and among paragraphs (or other meaningful chunks)	Lacks unity in constituent parts (such as paragraphs); fails to create coherence among constituent parts	Uses insufficient unifying statements (e.g., thesis statements, topic sentences, headings, or forecasting statements); uses few effective connections (e.g., transitions, match cuts, and hyperlinks)	Uses some effective unifying claims, but a few are unclear; makes connections weakly or inconsistently, as when claims appear as random lists or when paragraphs' topics lack explicit ties to the thesis	States unifying claims with supporting points that relate clearly to the overall argument and employs an effective but mechanical scheme	Asserts and sustains a claim that develops progressively and adapts typical organizational schemes for the context, achieving substantive coherence	Asserts a sophisticated claim by incorporating diverse perspectives that are organized to achieve maximum coherence and momentum
Conventions Expectations for grammar, mechanics, style, citation, and genre	Involves errors that risk making the overall message distorted or incomprehensible	Involves a major pattern of errors	Involves some distracting errors	Meets expectations, with minor errors	Exceeds expectations in a virtually flawless manner	Manipulates expectations in ways that advance the argument
Design for Medium Features that use affordances to enhance factors such as comprehensibility and usability	Lacks the features necessary for the genre; neglects significant affordances, such as linking on the web; uses features that conflict with or ignore the argument	Omits some important features; involves distracting inconsistencies in features (e.g., type and headings); uses features that don't support argument	Uses features that support with argument, but some match imprecisely with content; involves minor omissions or inconsistencies	Supports the argument with features that are generally suited to genre and content	Promotes engagement and supports the argument with features that efficiently use affordances	Persuades with careful, seamless integration of features and content and with innovative use of affordances

Appendix D

Final rubric used in summative evaluation.

Scale	1. Basic	2. Beginning	3. Developing	4. Competent	5. Mature	6. Exemplary
Rhetorical Awareness of the task at hand. Response to assignment considering elements such as purpose, register, and context.	Ignores multiple aspects of the situation and does not fulfill the task	Ignores at least one aspect of the situation and thus compromises effectiveness	Attempts to respond to all aspects of the situation, but the attempt is insufficient or inappropriate	Addresses the situation in a complete but perfunctory or predictable way	Addresses the situation completely, with unexpected insight	Addresses the situation in a complete, sophisticated manner that could advance professional discourse on the topic
Cites source text (using direct quotes or paraphrases) throughout to justify and demonstrate assertions made about it	Few if any citations, citations are virtually never used to support argument, or citations are consistently used out of context.	Inadequate citations, citations frequently don't connect to argument, or are taken out of context.	Tries to use citations to support argument. Provides nearly adequate citations, mostly used in context, but citations add little to the argument beyond what is already written.	Sufficient, contextually accurate, use of source text to support argument. Citations add to the argument, but in a predictable way.	Sufficient, compelling, precise, contextually accurate, use of source text to support argument.	Provides insightful analysis and exposition of citations to support argument. Citations are sufficient without being excessive, and used in context.
Articulates a position critically assessing a central aspect of the source text.	Does not affirm a position, or positions.	Takes very vague positions, or positions that do not address the article.	Takes one or more clear positions addressing the article, but lacks unity, or lacks significance ("so what?").	Offers a unified, significant, position addressing the thesis of the article, but position is a common or predictable one.	Offers a unified, significant position addressing the thesis of the article and takes a distinctive position.	Offers an inventive, expert-like position addressing the thesis of the article.
Assesses source text in the validity/form of its reasoning, in its content, and its overall success as an essay. Provides specific strengths weaknesses.	Provides no little or no attempt at assessing the text itself.	Assesses vague or trivial aspects of the text that are not relevant to the text's argument, or not clear how those aspects relate to the writer's position.	Identifies/assesses aspects of text that may be too general or not essential to its argument. Relationship of those aspects to the writer's position sometimes unclear.	Identifies/assesses specific, if obvious, aspects of the text's argument and relates them to the writer's position, but in a predictable way.	Identifies/assesses specific, core aspects of the text's argument, shows how these support the writer's position. Provides some suggestions for improvement.	Expertly identifies/assesses specific, core aspects of the text's argument in a way that compellingly supports writer's position. Insightfully explains how it could be improved.
Rational line of argument, including use of coherence devices. Arguments are logically sound, with any assumptions noted.	Involves an unspecified or confusing argument	Attempts to argue for certain positions and cites some evidence, but arguments have no logical force.	Makes reasonable arguments, but makes mistakes in form or in content.	Valid, well founded arguments, assumptions identified and explained where needed.	Sophisticated line of reasoning. Valid, well founded arguments, assumptions identified and explained where needed.	Creative line of reasoning. Valid, well founded, compelling arguments. *Necessary* assumptions are explained and defended.
Organization structure and coherence, including elements such. as introductions and conclusions as well as logical connections within and among paragraphs (or other meaningful chunks)	Lacks unity in constituent parts (such as paragraphs); fails to create coherence among constituent parts	Uses insufficient unifying statements (e.g., thesis statements, topic sentences, headings, or forecasting statements); uses few effective connections (e.g., transitions, match cuts, and hyperlinks)	Uses some effective unifying claims, but a few are unclear; makes connections weakly or inconsistently, as when claims appear as random lists or when paragraphs' topics lack explicit ties to the thesis	States unifying claims with supporting points that relate clearly to the overall argument and employs an effective but mechanical scheme	Asserts and sustains a claim that develops progressively and adapts typical organizational schemes for the context, achieving substantive coherence	Asserts a sophisticated claim by incorporating diverse perspectives that are organized to achieve maximum coherence and momentum
Conventions expectations for grammar, mechanics, style, citation, and genre	Involves errors that risk making the overall message distorted or incomprehensible	Involves a major pattern of errors	Involves some distracting errors	Meets expectations, with minor errors	Exceeds expectations in a virtually flawless manner	Manipulates expectations in ways that advance the argument

Appendix E

Participant timelines for summative study, and key for event numbers.

Navigation

1. **FDA**: Go forward definite amount (how much)
2. **BDA**: Go back definite amount (how much)
3. **SF:** Skim forward (how much)
4. **SB:** Skim backward (how much)
5. **FL**: Flip back and forth between two locations (how far apart are they).
6. **CO**: Manually pinch-**co**llapse the document (how much).
7. **UCO**: Manually **u**n**co**llapse the document (how? By button, or by touching scroll panel).
8. **LA:** Touch link arrow button (on what, and what moves in response).
9. **N2:** Navigate to two places at once by using two navigation operations (which ones)
10. **CH:** Collapse everything except highlights.
11. **UCH**: Un-collpase everything except highlights.
12. **PW**: Pan the workspace (how far).
13. **ZW**: Zoom the workspace (how much, in or out).
14. **ZD:** Zoom document in/out (how much, which direction).
15. **ZM:** Use a magnifying glass (What was done with the magnifying glass, what area was zoomed)

Excerpting

16. **CE:** Create excerpt and put on another document or the workspace (where in the workspace or other document).
17. **ME**: Move an excerpt (from where to where).
18. **RE:** Resize an excerpt (how much).
19. **OE**: Rotate an excerpt (how much).
20. **DE:** Erase/delete an excerpt.
21. **CT:** Copy text into the final document (how much, from where).

Annotation

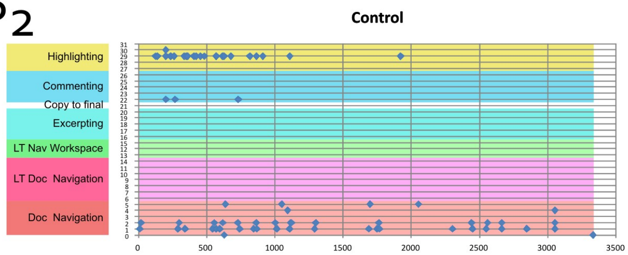
22. **CC:** Create comment (where is it located, what does it refer to, how long is it, is it near other marks).
23. **MC**:Move comment (from where to where).
24. **RC:** Resize comment (how much).
25. **OC:** Rotate comment (how much).
26. **DC:** Erase/delete comment.

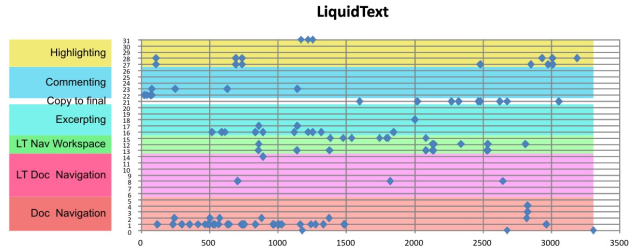
27. **H**: Highlight (what surface, using which color, is it near other marks)
28. **UH**: Unhighlight (on what surface).

29. **EM**: Draw emphasis marking (what mark, on what surface, is it near other marks).
30. **RM**: Draw relationship marking (what mark, relating what things, on what surface).
31. **LK**: Create a link in LiquidText (between what surface(s)).

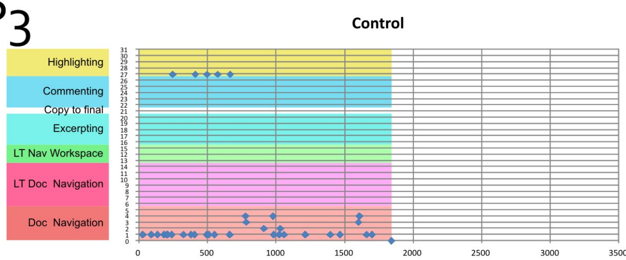
Timelines for participants: The x axis is time in seconds; end of colored area indicates task finished.
(Note that these charts are derived from the video data, which had several gaps due to capture errors.)

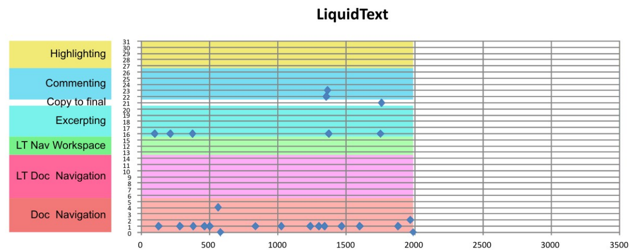
P2



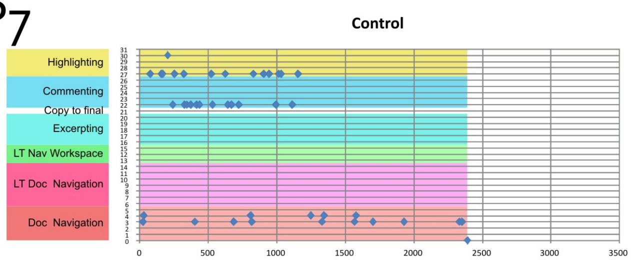


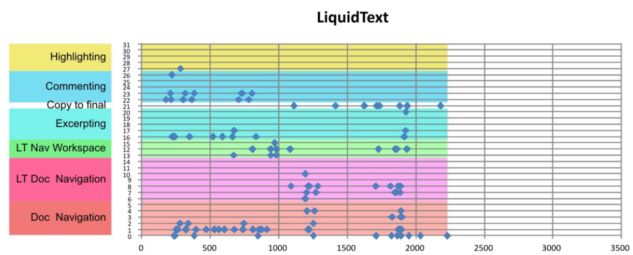
P3



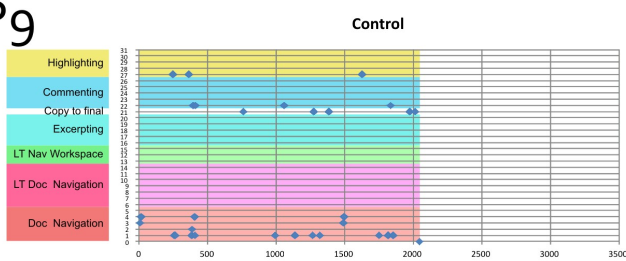


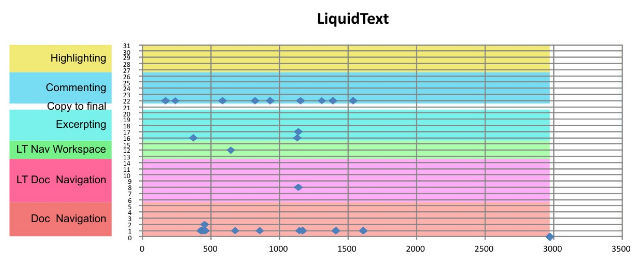
P7



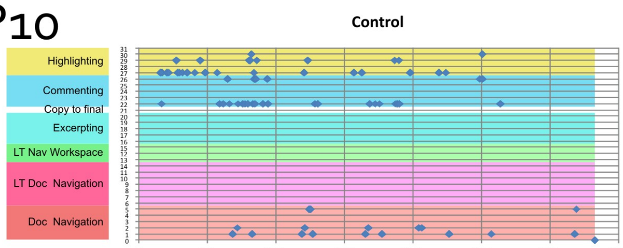


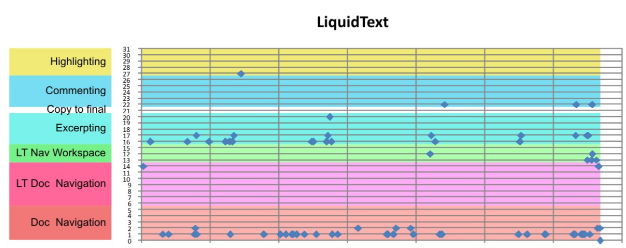
P9



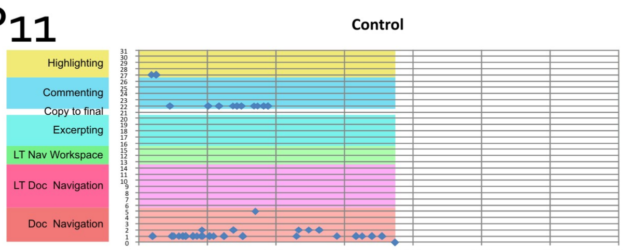


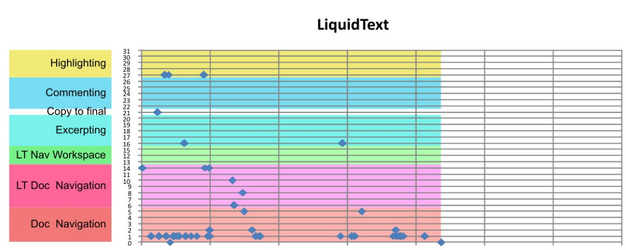
P10



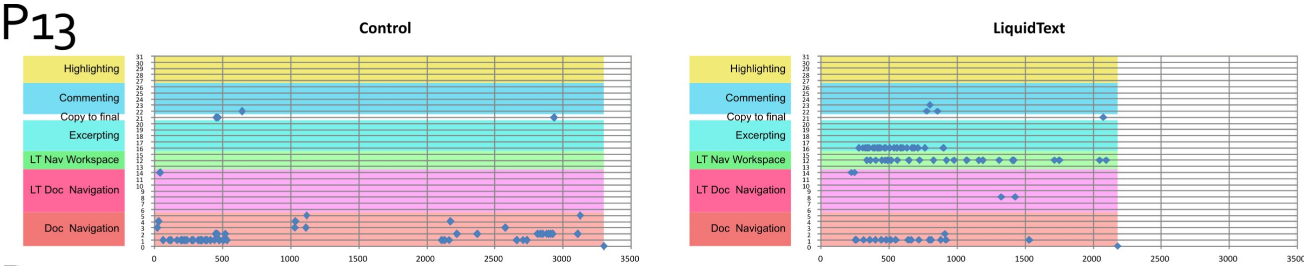


P11

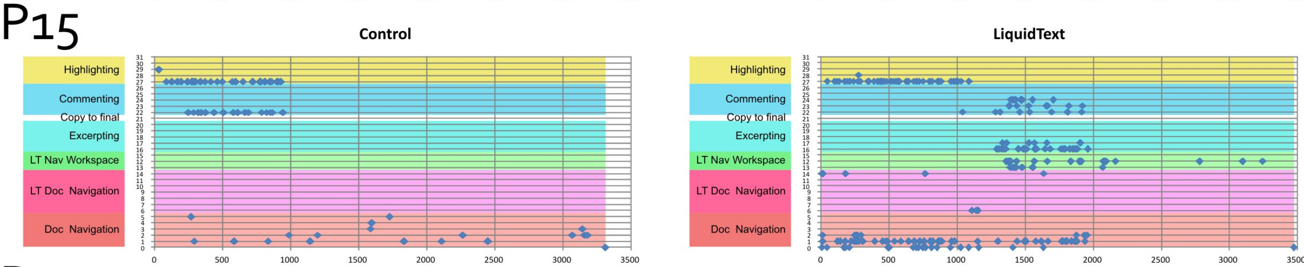




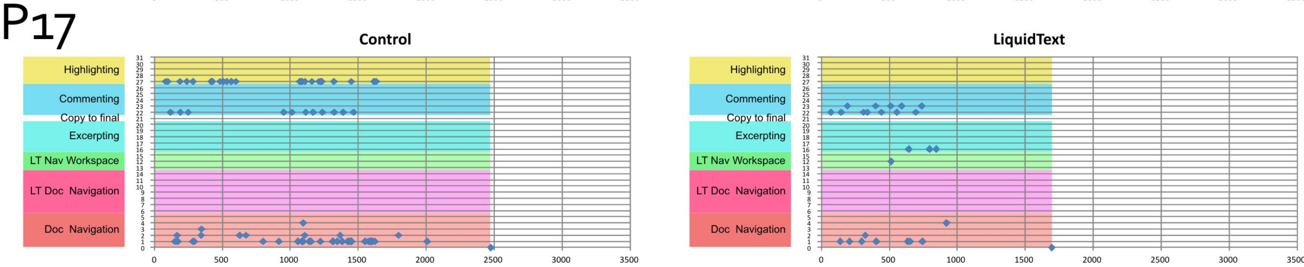
P13



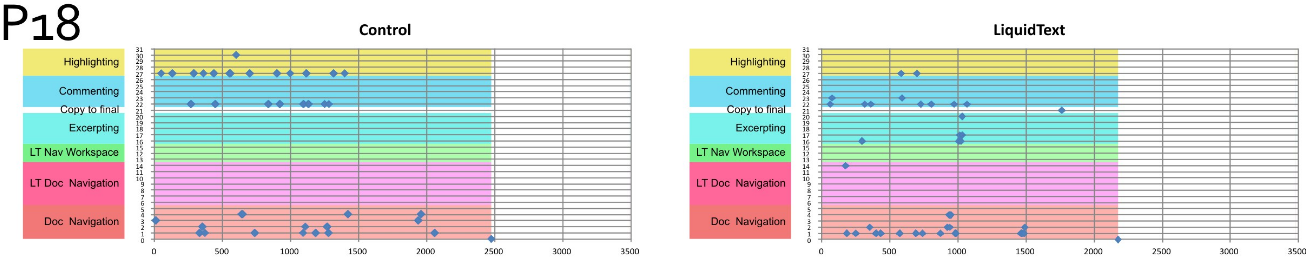
P15



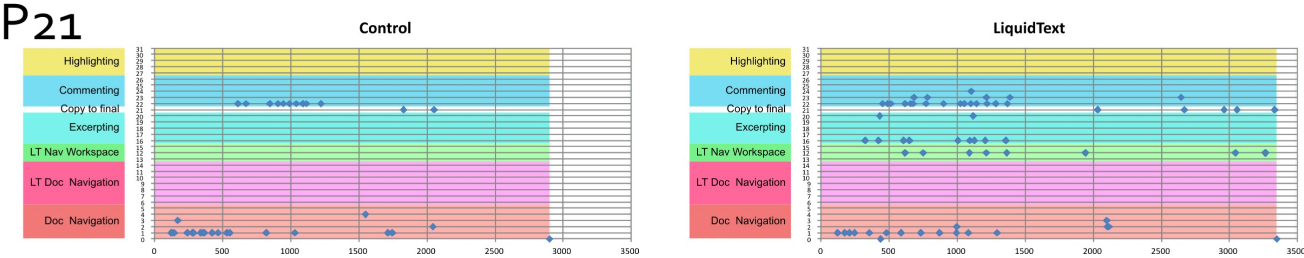
P17



P18



P21



References

Adler, A., Gujar, A., et al. (1998). A diary study of work-related reading: design implications for digital reading devices. CHI, ACM Press.

Adler, M. J. and VanDoren, C. (1972). How to Read a Book. New York, NY, Touchstone.

Apted, T., Sydney, A., et al. "Tabletop sharing of digital photographs for the elderly."

Araya, T., Akrami, N., et al. (2002). "Reducing Prejudice Through Priming of Control-Related Words." Experimental Psychology **49**(3): 222-227.

Askwall, S. (1985). "Computer supported reading vs. reading text on paper: a comparison of two reading situations." International Journal of Man Machine Studies **22**: 425-439.

Balakrishnan, R. and Hinckley, K. (2000). Symmetric bimanual interaction. Proceedings of the SIGCHI conference on Human factors in computing systems. The Hague, The Netherlands, ACM.

Baudisch, P. and Gutwin, C. (2004). Multiblending: displaying overlapping windows simultaneously without the drawbacks of alpha blending. CHI 2004, ACM Press.

Benko, H., Wilson, A. D., et al. (2006). Precise selection techniques for multi-touch screens, Google Patents.

Bernard, M. L., Chaparro, B. S., et al. (2003). "Comparing the effects of text size and format on the readibility of computer-displayed Times New Roman and Arial text." Int. J. Hum.-Comput. Stud. **59**(6): 823-835.

Bier, E. A., Stone, M. C., et al. (1994). Toolglass and magic lenses: the see-through interface. Conference companion on Human factors in computing systems. Boston, Massachusetts, United States, ACM Press.

Bordens, K. S. and Abbot, B. B. (2008). Research Design and Methods, Seventh Edition. Boston, McGraw Hill.

Bouvin, N. O., Zellweger, P. T., et al. (2002). Fluid annotations through open hypermedia: using and extending emerging web standards. Proc. World Wide Web. Honolulu, HI, USA, ACM.

Bredies, K., Mann, N. A., et al. (2008). The multi-touch SoundScape renderer. Proceedings of the working conference on Advanced visual interfaces. Napoli, Italy, ACM.

Broek, P. V. d., Young, M., et al. (1999). The Landscape Model of Reading: Inferences and he Online Construction of Memory Representation. The Construction of Mental Representations During Reading. H. vanOostendorp and S. R. Goldman. Mahwah, NJ, Laurence Erlbaum Associates.

Casalta, D., Guiard, Y., et al. (1999). Evaluating two-handed input techniques: rectangle editing and navigation. CHI '99 extended abstracts on Human factors in computing systems. Pittsburgh, Pennsylvania, ACM.

Catlin, T., Bush, P., et al. (1989). InterNote: extending a hypermedia framework to support annotative collaboration. Proceedings of the second annual ACM conference on Hypertext. Pittsburgh, Pennsylvania, United States, ACM.

Chang, B.-W., Mackinlay, J. D., et al. (1998). A negotiation architecture for fluid documents. Proceedings of the 11th annual ACM symposium on User interface software and technology. San Francisco, California, United States, ACM.

Costabile, M. F., Paternò, F., et al. (2005). Effects of Display Configurations on Document Triage. Human-Computer Interaction - INTERACT 2005, Springer Berlin / Heidelberg. **3585:** 130-143.

Davidson, P. L. and Han, J. Y. (2006). Synthesis and control on large scale multi-touch sensing displays. Proceedings of the 2006 conference on New interfaces for musical expression. Paris, France, IRCAM; Centre Pompidou.

Davidson, P. L. and Han, J. Y. (2008). Extending 2D object arrangement with pressure-sensitive layering cues. Proceedings of the 21st annual ACM symposium on User interface software and technology. Monterey, CA, USA, ACM.

Dietz, P. and Leigh, D. (2001). DiamondTouch: a multi-user touch technology. Proceedings of the 14th annual ACM symposium on User interface software and technology. Orlando, Florida, ACM.

Dillon, A., Richardson, J., et al. (1989). "Human factors of journal usage and design of electronic texts." Interacting with computers **1**: 183-189.

Egan, D. E., Remde, J. R., et al. (1989). "Formative design evaluation of superbook." ACM Trans. Inf. Syst. **7**(1): 30-57.

Eick, S. G., Steffen, J. L., et al. (1992). "Seesoft-A tool for Visualizing Line Oriented Software Statistics." IEEE Transactions on Software Engineering **18**: 957-968.

Elrod, S., Bruce, R., et al. (1992). Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. Proceedings of the SIGCHI conference on Human factors in computing systems. Monterey, California, United States, ACM.

Epps, J., Lichman, S., et al. (2006). A study of hand shape use in tabletop gesture interaction. CHI '06 extended abstracts on Human factors in computing systems. Montr\&\#233;al, Qu\&\#233;bec, Canada, ACM.

Epps, J., Lichman, S., et al. (2006). A study of hand shape use in tabletop gesture interaction, ACM.

Esenther, A. and Ryall, K. (2006). Fluid DTMouse: better mouse support for touch-based interactions. Proceedings of the working conference on Advanced visual interfaces. Venezia, Italy, ACM.

Evans, B. (1989). Method of and apparatus for sensing the location, such as coordinates, of designated points on an electrically sensitive touch-screen surface, Google Patents.

Foltz, P. W. (1996). Comprehension, coherence, and strategies in hypertext and linear text. Comprehension, coherence, and strategies in hypertext and linear text. Rouet, Levonen, Dillon and Spiro. Hillsdale, NJ, Laurence Erlbaum.

Gallardo, D., Julia, C. F., et al. (2008). TurTan: a Tangible Programming Language for Creative Exploration.

Gingold, Y. I., Davidson, P. L., et al. (2006). A direct texture placement and editing interface. Proceedings of the 19th annual ACM symposium on User interface software and technology. Montreux, Switzerland, ACM Press.

Ginsburg, A., Marks, J., et al. (1996). A viewer for PostScript documents. Proceedings of the 9th annual ACM symposium on User interface software and technology. Seattle, Washington, United States, ACM.

Golovchinsky, G. (2008). Reading in the office. Proceeding of the 2008 ACM workshop on Research advances in large digital book repositories. Napa Valley, CA, USA, ACM.

Golovchinsky, G. and Denoue, L. (2002). Moving markup: repositioning freeform annotations. Proceedings of the 15th annual ACM symposium on User interface software and technology. Paris, France, ACM.

Gough, P. B. and Tunmer, W. E. (1986). "Decoding, Reading, and Reading Disability." Remedial and Special Education **7**(1): 6-10.

Greenberg, S. and Gutwin, C. (1996). Awareness Through Fisheye Views in Relaxed-WYSIWIS Groupware. Proceedings of Graphics Interface, Morgan-Kaufmann.

Grossman, T., Wigdor, D., et al. (2005). Multi-finger gestural interaction with 3D volumetric displays. ACM SIGGRAPH 2005 Papers. Los Angeles, California, ACM Press.

Guiard, Y. (1987). "Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model." Journal of Motor Behavior **19**(4).

Guimbretiere, F. (2003). Paper augmented digital documents. Proceedings of UIST'03, New York, ACM.

Han, J. Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. Proceedings UIST 2005. Seattle, WA, USA, ACM.

Hill, W. C., Hollan, J. D., et al. (1992). Edit wear and read wear. Proceedings of the SIGCHI conference on Human factors in computing systems. Monterey, California, United States, ACM.

Hinckley, K., Yatani, K., et al. Pen + touch = new tools. UIST 2010, ACM.

Hinrichs, U., Hancock, M., et al. (2007). Examination of Text-Entry Methods for Tabletop Displays. Horizontal Interactive Human-Computer Systems, 2007. TABLETOP '07. Second Annual IEEE International Workshop on.

Hinrichs, U., Schmidt, H., et al. Bubbletype: Enabling text entry within a walk-up tabletop installation, Technical report, University of Calgary, 2008.

Hirche, J., Bomark, P., et al. "Adaptive interface for text input on large-scale interactive surfaces." Proc. TABLETOP **8**: 163–166.

Holmquist, L. E. (1997). Focus+context visualization with flip zooming and the zoom browser. CHI '97 extended abstracts on Human factors in computing systems: looking to the future. Atlanta, Georgia, ACM.

Hornbaek, K. and Frokjaer, E. (2003). "Reading patterns and usability in visualizations of electronic documents." ACM Trans. Comput.-Hum. Interact. **10**(2): 119-149.

Huey, E. B. (1908). The psychology and pedagogy of reading. New York, Macmillan.

Izadi, S., Hodges, S., et al. (2007). ThinSight: integrated optical multi-touch sensing through thin form-factor displays. Proceedings of the 2007 workshop on Emerging displays technologies: images and beyond: the future of displays and interacton. San Diego, California, ACM.

Jakobsen, M. R. and Hornbaek, K. (2006). Evaluating a fisheye view of source code. Proceedings CHI 2006, ACM.

Jiang, H., Wigdor, D., et al. (2008). "System Design for the WeSpace: Linking Personal Devices to a Table-Centered Multi-User, Multi-Surface Environment." Proc. Tabletop 2008: 105-112.

Jones, J. A., Harrold, M. J., et al. (2001). Visualization for Fault Localization. International Conference on Software Engineering. Toronto, Ontario, Canada.

Kabbash, P., Buxton, W., et al. (1994). Two-handed input in a compound task. Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence. Boston, Massachusetts, United States, ACM.

Knez, I. (1995). "Effects of Indoor Lighting on Mood and Cognition." Journal of Environmental Psychology **15**: 39-51.

Koike, H., Sato, Y., et al. (2000). Interactive textbook and interactive Venn diagram: natural and intuitive interfaces on augmented desk system. Proceedings of the SIGCHI conference on Human factors in computing systems. The Hague, The Netherlands, ACM.

Kos, R. (1991). "Persistence of Reading Disabilities: The Voices of Four Middle School Students." American Education Research Journal **28**(4): 875-895.

Lee, M. K., Kiesler, S., et al. (2011). Mining Behavioral Economics to Design Persuasive Technology for Healthy Choices. CHI '11, ACM.

Lee, S. K., William, B., et al. (1985). "A multi-touch three dimensional touch-sensitive tablet." SIGCHI Bull. **16**(4): 21-25.

Lemur (2009). Jazzmutant.

Liao, C., Guimbretiere, F., et al. (2008). "Papiercraft: A gesture-based command system for interactive paper." ACM Trans. Comput.-Hum. Interact. **14**(4): 1-27.

Lorch, R. F., Lorch, E. P., et al. (1993). "College Students' Conditional Knowledge about Reading." Journal of Education Psychology **85**.

Loviscach, J. (2007). Two-finger input with a standard touch screen. Proceedings of the 20th annual ACM symposium on User interface software and technology. Newport, Rhode Island, USA, ACM.

Mackay, W. and Pagani, D. (1994). Video mosaic: laying out time in a physical space. Proceedings of the second ACM international conference on Multimedia. San Francisco, California, United States, ACM.

Marshall, C. C. (1997). Annotation: from paper books to the digital library. Proceedings of the second ACM international conference on Digital libraries. Philadelphia, Pennsylvania, United States, ACM.

Marshall, C. C. and Bly, S. (2005). Saving and using encountered information: implications for electronic periodicals. Proceedings CHI 2005. Portland, Oregon, USA, ACM.

Marshall, C. C. and Bly, S. (2005). Turning the page on navigation. Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries. Denver, CO, USA, ACM.

Marshall, C. C. and Brush, A. J. B. (2004). Exploring the relationship between personal and public annotations. Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries. Tuscon, AZ, USA, ACM.

Marshall, C. C., Price, M. N., et al. (1999). Introducing a digital library reading appliance into a reading group. Proc. of conf. Digital libraries. Berkeley, CA, US, ACM.

Marshall, C. C., Price, M. N., et al. (2001). Designing e-books for legal research. Proc. of ACM/IEEE conf. on Digital libraries. Roanoke, VA, USA, ACM.

Masoodian, M., McKoy, S., et al. (2007). Hands-on sharing: collaborative document manipulation on a tabletop display using bare hands. Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI. Hamilton, New Zealand, ACM.

Meyer, B. J. F., Brandt, D. M., et al. (1980). "Use of Top-Level Structure in Text: Key for Reading Comprehension of Ninth-Grade Students." Reading Research Quarterly **16**(1): 72-103.

Morris, M. R., Brush, A. J. B., et al. (2007). Reading Revisited: Evaluating the Usability of Digital Display Surfaces for Active Reading Tasks. Tabletop.

Motamedi, N. (2008). Hd touch: multi-touch and object sensing on a high definition lcd tv. CHI '08 extended abstracts on Human factors in computing systems. Florence, Italy, ACM.

Murray, T. (2003). Applying Text Comprehension and Active Reading Principles to Adaptive Hyperbooks. Cognitive Science, Boston, MA.

Muter, P., Latremouille, S. A., et al. (1982). "Extended reading of continuous text on television screens." Human Factors **24**: 501-508.

Nist, S. L. and Kirby, K. (1989). "The text marking patterns of college students." Reading Psychology **10**: 321-338.

Norman, D. (1988). The design of everyday things, Basic Books.

O'Hara, K. (1996). Towards a Typology of Reading Goals. RXRC Affordances of Paper Project. Cambridge, UK, Rank Xerox Research Center.

O'Hara, K. and Sellen, A. (1997). A comparison of reading paper and on-line documents. CHI 1997, ACM.

O'Hara, K., Smith, F., et al. (1998). Student readers' use of library documents: implications for library technologies. CHI 1998, ACM Press.

Pearson, J., Buchanan, G., et al. (2009). Improving annotations in digital documents. European conference on Research and advanced technology for digital libraries. Corfu, Greece, Springer-Verlag.

Perfetti, C. A., Rouet, J.-F., et al. (1999). Towards a Theory of Documents Representation. The Construction of Mental Representations During Reading. H. vanOostendorp and S. R. Goldman. MahWah, NJ, Laurence Erlbaum Associates.

Piper, A. M. and Hollan, J. D. (2009). Tabletop displays for small group study: affordances of paper and digital materials. Proceedings of the 27th international conference on Human factors in computing systems. Boston, MA, USA, ACM.

Price, M. N., Golovchinsky, G., et al. (1998). Linking by inking: trailblazing in a paper-like hypertext. Proc. of Conf. Hypertext and hypermedia, ACM.

Qayyum, A. (2005). Student readers as authors: Analyzing markings made on e-documents in private or shared reading environments. Association for Library and Information Sciences Education Annual Conference, Boston, MA.

Qayyum, M. A. (2008). "Capturing the online academic reading process." Inf. Process. Manage. **44**(2): 581-595.

Read, J. C., Gregory, P., et al. (2002). "An investigation of participatory design with children-informant, balanced and facilitated design." Proceedings of Interaction Design and Children 2002: 51-58.

Reiss, S. P. (2006). Visualizing program execution using user abstractions. Proceedings of the 2006 ACM symposium on Software visualization. Brighton, United Kingdom, ACM.

Renear, A., DeRose, S., et al. (1999). An Outline for a Functional Taxonomy of Annotation. Redmond, WA, Presented at Microsoft Research.

Rick, J., Harris, A., et al. (2009). Children designing together on a multi-touch tabletop: an analysis of spatial orientation and user interactions. Proceedings of the 8th International Conference on Interaction Design and Children. Como, Italy, ACM.

Ringel, M., Ryall, K., et al. (2004). Release, relocate, reorient, resize: fluid techniques for document sharing on multi-user interactive tables. CHI '04 extended abstracts on Human factors in computing systems. Vienna, Austria, ACM.

Roast, C., Ritchie, I., et al. (2002). "Re-creating the reader: supporting active reading in literary research." Commun. ACM **45**(10): 109-111.

Robertson, G. G. and Mackinlay, J. D. (1993). The document lens. Proceedings of the 6th annual ACM symposium on User interface software and technology. Atlanta, Georgia, United States, ACM.

Roediger, H. L. and Karpicke, J. D. (2006). "The power of testing memory: Basic research and implications for educational practice. Perspectives " Perspectives on Psychological Science **1**: 181-210.

Samuels, S. J. and Kamil, M. L. (2002). 7 Models of the Reading Process. Handbook of reading research. P. D. Pearson, Laurence Erlbaum Associates.

Schilit, B. N., Golovchinsky, G., et al. (1998). Beyond paper: supporting active reading with free form digital ink annotations. CHI 1998, ACM.

Schiphorst, T., Motamedi, N., et al. (2007). Applying an Aesthetic Framework of Touch for Table-Top Interactions.

Sears, A. (1991). "Improving Touchscreen Keyboards: Design issues and a comparison with other devices." Computers **3**: 253-269.

Sellen, A. and Harper, R. (1997). Paper as an analytic resource for the design of new technologies. CHI 1997, ACM.

Shipman, F. M., Marshall, C. C., et al. (1999). Beyond location: hypertext workspaces and non-linear views. Proceedings of the tenth ACM Conference on Hypertext and hypermedia : returning to our diverse roots: returning to our diverse roots. Darmstadt, Germany, ACM.

Sun_Microsystems (1992). Starfire.

Vandoren, P., Claesen, L., et al. (2009). FluidPaint: an Interactive Digital Painting System using Real Wet Brushes Interactive Tabletops and Surfaces, Banff, Canada, IEEE.

Wang, X. and Maurer, F. (2008). Tabletop AgilePlanner: A Tabletop-Based Project Planning Tool for Agile Software Development Teams. IEEE Tabletops and Interactive Surfaces, Amsterdam, the Netherlands, IEEE.

Wellner, P. (1993). "Interacting with paper on the DigitalDesk." Commun. ACM **36**(7): 87-96.

Wilson, A. D. (2004). TouchLight: an imaging touch screen and display for gesture-based interaction. Proceedings of the 6th international conference on Multimodal interfaces. State College, PA, USA, ACM.

Wilson, A. D. (2005). TouchLight: an imaging touch screen and display for gesture-based interaction. ACM SIGGRAPH 2005 Emerging technologies. Los Angeles, California, ACM.

Wobbrock, J. O., Morris, M. R., et al. (2009). User-defined gestures for surface computing. Proceedings of the 27th international conference on Human factors in computing systems. Boston, MA, USA, ACM.

Wolfe, J. L. (2000). Effects of annotations on student readers and writers. Proceedings of the fifth ACM conference on Digital libraries. San Antonio, Texas, United States, ACM.

Wu, M. and Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. UIST 2003, ACM.

Zeleznik, R., Bragdon, A., et al. (2010). Hands-on math: a page-based multi-touch and pen desktop for technical work and problem solving. Proc of ACM UIST. New York, NY, USA, ACM.

Zellweger, P. T., Regli, S. H., et al. (2000). The impact of fluid documents on reading and browsing: an observational study. Proceedings of the SIGCHI conference on Human factors in computing systems. The Hague, The Netherlands, ACM.

Ziemkiewicz, C. and Kosara, R. Implied dynamics in information visualization. Proceedings of the International Conference on Advanced Visual Interfaces. Roma, Italy, ACM.